518 Rec'd PCT/PTO 31 JUL 2001

FORM PTO-1390 U S DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371		ATTORNEY'S DOCKET NUMBER **PF-0676 USN** U.S. APPLICATION NO. (If known, see 37 CFR 1.5) **TO BE ASSIGNED** 09/890549
INTERNATIONAL APPLICATION NO. **PCT/US00/04160**	INTERNATIONAL FILING DATE **18 February 2000**	PRIORITY DATE CLAIMED **19 February 1999**

TITLE OF INVENTION
HUMAN LIPID ASSOCIATED PROTEINS

APPLICANT(S) FOR DO/EO/US
INCYTE PHARMACEUTICALS, INC.; TANG, Y. Tom; HILLMAN, Jennifer L.; YUE, Henry; AZIMZAI, Yalda; BAUGHN, Mariah R.; TRAN, Bao

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. ☒ This is the **FIRST** submission of items concerning a filing under 35 U.S.C. 371.
2. ☐ This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371.
3. ☐ This is an express request to promptly begin national examination procedures (35 U.S.C. 371 (f)).
4. ☐ The US has been elected by the expiration of 19 months from the priority date (PCT Article 31).
5. ☒ A copy of the International Application as filed (35 U.S.C. 371(c)(2))
   a. ☐ is attached hereto (required only if not communicated by the International Bureau)
   b. ☐ has been communicated by the International Bureau.
   c. ☒ is not required, as the application was filed in the United States Receiving Office (RO/US).
6. ☐ An English language translation of the International Application as filed (35 U.S.C. 371(c)(2)).
7. ☒ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))
   a. ☐ are attached hereto (required only if not communicated by the International Bureau).
   b. ☐ have been communicated by the International Bureau.
   c. ☐ have not been made; however, the time limit for making such amendments has NOT expired.
   d. ☒ have not been made and will not be made.
8. ☐ An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).
9. ☐ An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).
10. ☐ An English language translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).

Items 11 to 16 below concern document(s) or information included:

11. ☐ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
12. ☐ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.27 and 3.31 is included.
13. ☐ A FIRST preliminary amendment.
    ☐ A SECOND or SUBSEQUENT preliminary amendment.
14. ☐ A substitute specification.
15. ☐ A change of power of attorney and/or address letter.
16. ☒ Other items or information:

1) Transmittal Letter (2 pp, in duplicate)
2) Return Postcard
3) Express Mail Label No.: **EL 856 154 177 US**
4) Request to Transfer

JC05 Rec'd PCT/PTO 31 JUL 2001

U.S. APPLICATION NO. (if known, see 37 CFR 1.5) TO BE ASSIGNED 09/890549	INTERNATIONAL APPLICATION NO.: PCT/US00/04160	ATTORNEY'S DOCKET NUMBER PF-0676 USN

17. ☒ The following fees are submitted:
BASIC NATIONAL FEE (37 CFR 1.492(a)(1)-(5):
Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO.....$1000.00
☐ International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO..$860.00
International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee (37 CFR 1.445(a)(2)) paid to USPTO.....$710.00
☒ International preliminary examination fee paid to USPTO (37 CFR 1.482) but all claims did not satisfy provisions of PCT Article 33(1)-(4).....$690.00
☐ International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(1)-(4).....$100.00

ENTER APPROPRIATE BASIC FEE AMOUNT =				$690.00	
Surcharge of **$130.00** for furnishing the oath or declaration later than ☐ 20 ☐ 30 months from the earliest claimed priority date (37 CFR 1.492(e)).				$	
CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE		
Total Claims	23 =	3	X **$ 18.00**	$ 54.00	
Independent Claims	2 =	0	X **$ 80.00**	$	
MULTIPLE DEPENDENT CLAIM(S) (if applicable)			+ **$270.00**	$	
TOTAL OF ABOVE CALCULATIONS =				$744.00	
☐ Applicant claims small entity status See 37 CFR 1.27. The fees indicated above are reduced by 1/2.				$	
SUBTOTAL =				$744.00	
Processing fee of **$130.00** for furnishing the English translation later than ☐ 20 ☐ 30 months from the earliest claimed priority date (37 CFR 1.492(f)). +				$	
TOTAL NATIONAL FEE =				$744.00	
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by the appropriate cover sheet (37 CFR 3.28, 3.31). **$40.00** per property +				$	
TOTAL FEES ENCLOSED =				$744.00	
				Amount to be Refunded:	$
				Charged	$

a ☐ A check in the amount of $_____ to cover the above fees is enclosed.
b. ☒ Please charge my Deposit Account No. 09-0108 in the amount of $ 744.00 to cover the above fees.
c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 09-0108. A duplicate copy of this sheet is enclosed.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

INCYTE GENOMICS, INC
3160 Porter Drive
Palo Alto, CA 94304

Diana Hamlet-Cox
SIGNATURE

NAME: Diana Hamlet-Cox

REGISTRATION NUMBER: 33,302

DATE: 31 July 2001

HUMAN LIPID-ASSOCIATED PROTEINS

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of human lipid-associated proteins and to the use of these sequences in the diagnosis, treatment, and prevention of cardiovascular, neurological, and gastrointestinal disorders, and disorders of lipid metabolism.

BACKGROUND OF THE INVENTION

Lipids are water-insoluble, oily, or greasy substances that are soluble in nonpolar solvents such as chloroform or ether. Neutral lipids (triacylglycerols) serve as major fuels and energy stores. Polar lipids, such as phospholipids, sphingolipids, glycolipids, and cholesterol, are key structural components of cell membranes. Lipids and proteins are associated in a varicty of ways. Glycolipids form vesicles that carry proteins within cells and cell membranes. Interactions between lipids and proteins function in targeting proteins and glycolipids involved in a variety of processes, such as cell signaling and cell proliferation, to specific membrane and intracellular locations. Proteins are associated with the biosynthesis, transport, and uptake of lipids. In addition, key proteins involved in signal transduction and protein targeting have lipid-derived groups added to them post-translationally (Stryer, L. (1995) Biochemistry, W.H. Freeman and Co., New York NY, pp. 264-267, 934).

A major class of phospholipids are the phosphoglycerides, which are composed of a glycerol backbone, two fatty acid chains, and a phosphorylated alcohol. Principal phosphoglycerides are phosphatidylcholine, phosphatidylethanolamine, phosphatidylserine, phosphatidylinositol, and diphosphatidylglycerol. Many enzymes involved in phosphoglyceride synthesis are associated with membranes (Meyers, R.A. (1995) Molecular Biology and Biotechnology, VCH Publishers Inc., New York NY, pp. 494-501; Stryer, supra, pp. 264-267). The enzyme phosphatidylserine decarboxylase catalyzes the conversion of phosphatidylserine to phosphatidylethanolamine, using a pyruvate cofactor. The two forms of yeast phosphatidylserine decarboxylase are localized to the inner mitochondrial membrane and to the Golgi/vacuole membrane, respectively. The mammalian enzyme, also localized to the inner mitochondrial membrane, is made as a proenzyme and subsequently cleaved to alpha and beta subunits (Voelker, D.R. (1997) Biochim. Biophys. Acta 1348:236-244).

Cholesterol, composed of four fused hydrocarbon rings with an alcohol at one end, moderates the fluidity of membranes in which it is incorporated. In addition, cholesterol is used in the synthesis of such hormones as cortisol, progesterone, estrogen, and testosterone. Bile salts derived from cholesterol facilitate the digestion of lipids. Cholesterol in the skin forms a barrier that prevents excess water evaporation from the body. Farnesyl and geranylgeranyl groups, which are derived from cholesterol biosynthesis intermediates, are post-translationally added to signal transduction proteins

such as ras and protein-targeting proteins such as rab. These modifications are important for the activities of these proteins (Guyton, A.C. Textbook of Medical Physiology (1991) W.B. Saunders Company, Philadelphia PA, pp.760-763; Stryer, supra, pp. 279-280, 691-702, 934).

Mammals obtain cholesterol derived from both de novo biosynthesis and the diet. The liver is the major site of cholesterol biosynthesis in mammals. Biosynthesis is accomplished via a series of enzymatic steps known as the mevalonate pathway. The rate-limiting step is the conversion of hydroxymethylglutaryl-Coenzyme A (HMG-CoA) to mevalonate by HMG-CoA reductase. The drug lovastatin, a potent inhibitor of HMG-CoA reductase, is given to patients to reduce their serum cholesterol levels. Cholesterol derived from de novo biosynthesis or from the diet is transported in the body fluids in the form of lipoprotein particles. These particles also transport triacylglycerols. The particles consist of a core of hydrophobic lipids surrounded by a shell of polar lipids and apolipoproteins. The protein components serve in the solubilization of hydrophobic lipids and also contain cell-targeting signals. Lipoproteins include chylomicrons, chylomicron remnants, very-low-density lipoproteins (VLDL), intermediate-density lipoproteins (IDL), low-density lipoproteins (LDL), and high-density lipoproteins (HDL) (Meyers, supra; Stryer, supra, pp. 691-702). There is a strong inverse correlation between the levels of plasma HDL and risk of premature coronary heart disease. ApoL is an HDL apolipoprotein expressed in the pancreas (Duchateau, P.N. et al. (1997) J. Biol. Chem. 272:25576-25582).

Most cells outside the liver and intestine take up cholesterol from the blood rather than synthesize it themselves. Cell surface LDL receptors bind LDL particles which are then internalized by endocytosis (Meyers, supra). Absence of the LDL receptor, the cause of the disease familial hypercholesterolemia, leads to increased plasma cholesterol levels and ultimately to atherosclerosis (Stryer, supra, pp. 691-702).

Proteins involved in cholesterol uptake and biosynthesis are tightly regulated in response to cellular cholesterol levels. The sterol regulatory element binding protein (SREBP) is a sterol-responsive transcription factor. Under normal cholesterol conditions, SREBP resides in the endoplasmic reticulum membrane. When cholesterol levels are low, a regulated cleavage of SREBP occurs which releases the extracellular domain of the protein. This cleaved domain is then transported to the nucleus where it activates the transcription of the LDL receptor gene, and genes encoding enzymes of cholesterol synthesis, by binding the sterol regulatory element (SRE) upstream of the genes (Yang, J. et al. (1995) J. Biol. Chem. 270:12152-12161). Regulation of cholesterol uptake and biosynthesis also occurs via the oxysterol-binding protein (OSBP). Oxysterols are oxidation products formed during the catabolism of cholesterol, and are involved in regulation of steroid biosynthesis. OSBP is a high-affinity intracellular receptor for a variety of oxysterols that down-regulate cholesterol synthesis and stimulate cholesterol esterification (Lagace, T.A. et al. (1997) Biochem. J. 326:205-213).

The copines are phospholipid-binding proteins believed to function in membrane trafficking. Copines promote lipid vesicle aggregation. They contain a C2 domain associated with membrane activity and an annexin-type domain that mediates interactions between integral and extracellular proteins and is associated with calcium binding and regulation (Creutz, C.E. (1998) J. Biol. Chem. 273:1393-1402). Other C2-containing proteins include the synaptotagmins, a family of proteins involved in vesicular trafficking. Synaptotagmin concentrations in cerebrospinal fluid have been found to be reduced in early-onset Alzheimer's disease (Gottfries, C.G. et al. (1998) J. Neural Transm. 105:773-786).

Lipids and their associated proteins have roles in human diseases and disorders. Increased synthesis of long-chain fatty acids occurs in neoplasms including those of the breast, prostate, ovary, colon and endometrium. There is a strong inverse correlation between the levels of plasma HDL and risk of premature coronary heart disease. Absence of the LDL receptor, the cause of familial hypercholesterolemia, leads to increased plasma cholesterol levels and ultimately to atherosclerosis (Stryer, supra, pp. 691-702). The arterial disease atherosclerosis is characterized by the formation of fatty lesions on the inside of the arterial wall. These lesions promote the loss of arterial flexibility and the formation of blood clots (Guyton, supra). Oxysterols are present in human atherosclerotic plaques and believed to play an active role in plaque development (Brown, A.J. (1999) Atherosclerosis 142:1-28). Steatosis, or fatty liver, is characterized by the accumulation of triglycerides in the liver and may occur in association with a variety of conditions including alcoholism, diabetes, obesity, and prolonged parenteral nutrition. Steatosis may lead to fibrosis and cirrhosis of the liver. In Tay-Sachs disease, the GM_2 ganglioside (a sphingolipid) accumulates in lysosomes of the central nervous system due to a lack of the enzyme N-acetylhexosaminidase. Patients suffer nervous system degeneration leading to early death (Fauci, A.S. et al. (1998) Harrison's Principles of Internal Medicine McGraw-Hill, New York NY p. 2171). The Niemann-Pick diseases are caused by defects in lipid metabolism. Niemann-Pick diseases types A and B are caused by accumulation of sphingomyelin (a sphingolipid) and other lipids in the central nervous system due to a defect in the enzyme sphingomyelinase, leading to neurodegeneration and lung disease. Niemann-Pick disease type C results from a defect in cholesterol transport, leading to the accumulation of sphingomyelin and cholesterol in lysosomes and a secondary reduction in sphingomyelinase activity. Neurological symptoms such as grand mal seizures, ataxia, and loss of previously learned speech, manifest 1-2 years after birth. A mutation in the NPC protein, which contains a putative cholesterol-sensing domain, was found in a mouse model of Niemann-Pick disease type C (Fauci, supra, p. 2175; Loftus, S.K. et al. (1997) Science 277:232-235).

The discovery of new human lipid-associated proteins and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of cardiovascular, neurological, and gastrointestinal disorders, and disorders of lipid

metabolism.

SUMMARY OF THE INVENTION

The invention features purified polypeptides, human lipid-associated proteins, referred to collectively as "LIPAP" and individually as "LIPAP-1," "LIPAP-2," "LIPAP-3," "LIPAP-4," "LIPAP-5," "LIPAP-6," "LIPAP-7," "LIPAP-8," "LIPAP-9," "LIPAP-10," "LIPAP-11," and "LIPAP-12." In one aspect, the invention provides an isolated polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-12. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-12.

The invention further provides an isolated polynucleotide encoding a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-12. In one alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:13-24.

Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-12. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

The invention also provides a method for producing a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-12. The method comprises a) culturing a

cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Additionally, the invention provides an isolated antibody which specifically binds to a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-12.

The invention further provides an isolated polynucleotide comprising a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:13-24, b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:13-24, c) a polynucleotide sequence complementary to a), or d) a polynucleotide sequence complementary to b). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide comprising a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:13-24, b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:13-24, c) a polynucleotide sequence complementary to a), or d) a polynucleotide sequence complementary to b). The method comprises a) hybridizing the sample with a probe comprising at least 16 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 30 contiguous nucleotides. In another alternative, the probe comprises at least 60 contiguous nucleotides.

The invention further provides a pharmaceutical composition comprising an effective amount of a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, and a pharmaceutically acceptable excipient. The invention additionally provides a method of

treating a disease or condition associated with decreased expression of functional LIPAP, comprising administering to a patient in need of such treatment the pharmaceutical composition.

The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-12. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a pharmaceutical composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with decreased expression of functional LIPAP, comprising administering to a patient in need of such treatment the pharmaceutical composition.

Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-12, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-12. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a pharmaceutical composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with overexpression of functional LIPAP, comprising administering to a patient in need of such treatment the pharmaceutical composition.

The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence selected from the group consisting of SEQ ID NO:13-24, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered expression of the target polynucleotide.

BRIEF DESCRIPTION OF THE TABLES

Table 1 shows polypeptide and nucleotide sequence identification numbers (SEQ ID NOs), clone identification numbers (clone IDs), cDNA libraries, and cDNA fragments used to assemble full-

length sequences encoding LIPAP.

Table 2 shows features of each polypeptide sequence, including potential motifs, homologous sequences, and methods, algorithms, and searchable databases used for analysis of LIPAP.

Table 3 shows selected fragments of each nucleic acid sequence; the tissue-specific expression patterns of each nucleic acid sequence as determined by northern analysis; diseases, disorders, or conditions associated with these tissues; and the vector into which each cDNA was cloned.

Table 4 describes the tissues used to construct the cDNA libraries from which cDNA clones encoding LIPAP were isolated.

Table 5 shows the tools, programs, and algorithms used to analyze LIPAP, along with applicable descriptions, references, and threshold parameters.

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"LIPAP" refers to the amino acid sequences of substantially purified LIPAP obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of

LIPAP. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of LIPAP either by directly interacting with LIPAP or by acting on components of the biological pathway in which LIPAP participates.

An "allelic variant" is an alternative form of the gene encoding LIPAP. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding LIPAP include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as LIPAP or a polypeptide with at least one functional characteristic of LIPAP. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding LIPAP, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding LIPAP. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent LIPAP. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of LIPAP is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

The terms "amino acid" and "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited to refer to an amino acid sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity

of LIPAP. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of LIPAP either by directly interacting with LIPAP or by acting on components of the biological pathway in which LIPAP participates.

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, $F(ab')_2$, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind LIPAP polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition capable of base-pairing with the "sense" strand of a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic LIPAP, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The terms "complementary" and "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence "5' A-G-T 3'" bonds to the complementary

sequence "3' T-C-A 5'." Complementarity between two single-stranded molecules may be "partial," such that only some of the nucleic acids bind, or it may be "complete," such that total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acid strands, and in the design and use of peptide nucleic acid (PNA) molecules.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding LIPAP or fragments of LIPAP may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been resequenced to resolve uncalled bases, extended using the XL-PCR kit (Perkin-Elmer, Norwalk CT) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from the overlapping sequences of one or more Incyte Clones and, in some cases, one or more public domain ESTs, using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that, when made, least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

Original Residue	Conservative Substitution
Ala	Gly, Ser
Arg	His, Lys
Asn	Asp, Gln, His
Asp	Asn, Glu
Cys	Ala, Ser
Gln	Asn, Glu, His
Glu	Asp, Gln, His
Gly	Ala
His	Asn, Arg, Gln, Glu
Ile	Leu, Val
Leu	Ile, Val
Lys	Arg, Gln, Glu
Met	Leu, Ile

Phe	His, Met, Leu, Trp, Tyr
Ser	Cys, Thr
Thr	Ser, Val
Trp	Phe, Tyr
Tyr	His, Phe, Trp
Val	Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to the chemical modification of a polypeptide sequence, or a polynucleotide sequence. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "fragment" is a unique portion of LIPAP or the polynucleotide encoding LIPAP which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50% of a polypeptide) as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:13-24 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:13-24, for example, as distinct from any other sequence in the same genome. A fragment of SEQ ID NO:13-24 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:13-24 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:13-24 and the region of SEQ ID NO:13-24 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based

on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-12 is encoded by a fragment of SEQ ID NO:13-24. A fragment of SEQ ID NO:1-12 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-12. For example, a fragment of SEQ ID NO:1-12 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-12. The precise length of a fragment of SEQ ID NO:1-12 and the region of SEQ ID NO:1-12 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

The term "similarity" refers to a degree of complementarity. There may be partial similarity or complete similarity. The word "identity" may substitute for the word "similarity." A partially complementary sequence that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid is referred to as "substantially similar." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or northern blot, solution hybridization, and the like) under conditions of reduced stringency. A substantially similar sequence or hybridization probe will compete for and inhibit the binding of a completely similar (identical) sequence to the target sequence under conditions of reduced stringency. This is not to say that conditions of reduced stringency are such that non-specific binding is permitted, as reduced stringency conditions require that the binding of two sequences to one another be a specific (i.e., a selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% similarity or identity). In the absence of non-specific binding, the substantially similar sequence or probe will not hybridize to the second non-complementary target sequence.

The phrases "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned

polynucleotide sequence pairs.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at http://www.ncbi.nlm.nih.gov/BLAST/. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at http://www.ncbi.nlm.nih.gov/gorf/bl2.html. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Reward for match: 1

Penalty for mismatch: -2

Open Gap: 5 and Extension Gap: 2 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 11

Filter: on

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a

standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the hydrophobicity and acidity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table. As with polynucleotide alignments, the percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) with blastp set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for stable mitotic chromosome segregation and maintenance.

The term "humanized antibody" refers to antibody molecules in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific

hybridization is an indication that two nucleic acid sequences share a high degree of identity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Generally, such wash temperatures are selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating T_m and conditions for nucleic acid hybridization are well known and can be found in Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, denatured salmon sperm DNA at about 100-200 µg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

The term "microarray" refers to an arrangement of distinct polynucleotides on a substrate.

The terms "element" and "array element" in a microarray context, refer to hybridizable polynucleotides arranged on the surface of a substrate.

The term "modulate" refers to a change in the activity of LIPAP. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of LIPAP.

The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with the second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Generally, operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Probe" refers to nucleic acid sequences encoding LIPAP, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous

nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for example Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel et al.,1987, Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis et al., 1990, PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence

that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, supra. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be use to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

An "RNA equivalent," in reference to a DNA sequence, is composed of the same linear sequence of nucleotides as the reference DNA sequence with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The term "sample" is used in its broadest sense. A sample suspected of containing nucleic acids encoding LIPAP, or fragments thereof, or LIPAP itself, may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide containing the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least 60% free, preferably at least 75% free, and most preferably at least 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

"Transformation" describes a process by which exogenous DNA enters and changes a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed" cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "transgenic organism," as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or in vitro fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, and plants and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook et al. (1989), supra.

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95% or at least 98% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide

polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 98% or greater sequence identity over a certain defined length of one of the polypeptides.

THE INVENTION

The invention is based on the discovery of new human lipid-associated proteins (LIPAP), the polynucleotides encoding LIPAP, and the use of these compositions for the diagnosis, treatment, or prevention of cardiovascular, neurological, and gastrointestinal disorders, and disorders of lipid metabolism.

Table 1 lists the Incyte clones used to assemble full length nucleotide sequences encoding LIPAP. Columns 1 and 2 show the sequence identification numbers (SEQ ID NOs) of the polypeptide and nucleotide sequences, respectively. Column 3 shows the clone IDs of the Incyte clones in which nucleic acids encoding each LIPAP were identified, and column 4 shows the cDNA libraries from which these clones were isolated. Column 5 shows Incyte clones and their corresponding cDNA libraries. Clones for which cDNA libraries are not indicated were derived from pooled cDNA libraries. The Incyte clones in column 5 were used to assemble the consensus nucleotide sequence of each LIPAP and are useful as fragments in hybridization technologies.

The columns of Table 2 show various properties of each of the polypeptides of the invention: column 1 references the SEQ ID NO; column 2 shows the number of amino acid residues in each polypeptide; column 3 shows potential phosphorylation sites; column 4 shows potential glycosylation sites; column 5 shows the amino acid residues comprising signature sequences and motifs; column 6 shows homologous sequences as identified by BLAST analysis and the identity of each polypeptide; and column 7 shows analytical methods and in some cases, searchable databases to which the analytical methods were applied. The methods of column 7 were used to characterize each polypeptide through sequence homology and protein motifs.

The columns of Table 3 show the tissue-specificity and diseases, disorders, or conditions associated with nucleotide sequences encoding LIPAP. The first column of Table 3 lists the nucleotide SEQ ID NOs. Column 2 lists fragments of the nucleotide sequences of column 1. These fragments are useful, for example, in hybridization or amplification technologies to identify SEQ ID NO:13-24 and to distinguish between SEQ ID NO:13-24 and related polynucleotide sequences. The polypeptides

encoded by these fragments are useful, for example, as immunogenic peptides. Column 3 lists tissue categories which express LIPAP as a fraction of total tissues expressing LIPAP. Column 4 lists diseases, disorders, or conditions associated with those tissues expressing LIPAP as a fraction of total tissues expressing LIPAP. Column 5 lists the vectors used to subclone each cDNA library.

The columns of Table 4 show descriptions of the tissues used to construct the cDNA libraries from which cDNA clones encoding LIPAP were isolated. Column 1 references the nucleotide SEQ ID NOs, column 2 shows the cDNA libraries from which these clones were isolated, and column 3 shows the tissue origins and other descriptive information relevant to the cDNA libraries in column 2.

SEQ ID NO:21 maps to chromosome 11 within the interval from 92.5 to 96.3 centiMorgans. This interval also contains a gene encoding a G-protein coupled receptor associated with epilepsy.

The invention also encompasses LIPAP variants. A preferred LIPAP variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the LIPAP amino acid sequence, and which contains at least one functional or structural characteristic of LIPAP.

The invention also encompasses polynucleotides which encode LIPAP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:13-24, which encodes LIPAP. The polynucleotide sequences of SEQ ID NO:13-24, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses a variant of a polynucleotide sequence encoding LIPAP. In particular, such a variant polynucleotide sequence will have at least about 80%, or alternatively at least about 90%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding LIPAP. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:13-24 which has at least about 80%, or alternatively at least about 90%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:13-24. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of LIPAP.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding LIPAP, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally

occurring LIPAP, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode LIPAP and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring LIPAP under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding LIPAP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding LIPAP and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode LIPAP and LIPAP derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding LIPAP or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:13-24 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Perkin-Elmer), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Perkin-Elmer). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Perkin-Elmer), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and

Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding LIPAP may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Perkin-Elmer), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode LIPAP may be cloned in recombinant DNA molecules that direct expression of LIPAP, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express LIPAP.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter LIPAP-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent Number 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Crameri, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of LIPAP, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

In another embodiment, sequences encoding LIPAP may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; and Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.) Alternatively, LIPAP itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solid-phase techniques. (See, e.g., Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Perkin-Elmer). Additionally, the amino acid sequence of LIPAP, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to

produce a variant polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY.)

In order to express a biologically active LIPAP, the nucleotide sequences encoding LIPAP or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding LIPAP. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding LIPAP. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding LIPAP and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding LIPAP and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding LIPAP. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding LIPAP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding LIPAP can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding LIPAP into the vector's multiple cloning site disrupts the *lac*Z gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of LIPAP are needed, e.g. for the production of antibodies, vectors which direct high level expression of LIPAP may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of LIPAP. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; and Scorer, C.A. et al. (1994) Bio/Technology 12:181-184.)

Plant systems may also be used for expression of LIPAP. Transcription of sequences encoding LIPAP may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding LIPAP may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses LIPAP in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-

based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of LIPAP in cell lines is preferred. For example, sequences encoding LIPAP can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk*$^-$ and *apr*$^-$ cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), ß glucuronidase and its substrate ß-glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding LIPAP is inserted within a marker gene sequence, transformed cells containing sequences encoding LIPAP can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding LIPAP under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates

expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding LIPAP and that express LIPAP may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of LIPAP using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on LIPAP is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding LIPAP include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding LIPAP, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding LIPAP may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode LIPAP may be designed to contain signal sequences which direct secretion of LIPAP through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the

polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding LIPAP may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric LIPAP protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of LIPAP activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the LIPAP encoding sequence and the heterologous protein sequence, so that LIPAP may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled LIPAP may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, ^{35}S-methionine.

Fragments of LIPAP may be produced not only by recombinant means, but also by direct peptide synthesis using solid-phase techniques. (See, e.g., Creighton, supra, pp. 55-60.) Protein synthesis may be performed by manual techniques or by automation. Automated synthesis may be achieved, for example, using the ABI 431A peptide synthesizer (Perkin-Elmer). Various fragments of LIPAP may be synthesized separately and then combined to produce the full length molecule.

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of LIPAP and human lipid-associated proteins. In addition, the expression of LIPAP is closely associated with cardiovascular and gastrointestinal tissues, and tissues of the nervous system. Therefore, LIPAP appears to play a role in cardiovascular, neurological, and gastrointestinal disorders, and disorders of lipid metabolism. In the treatment of disorders associated with increased LIPAP expression or activity, it is desirable to decrease the expression or activity of LIPAP. In the treatment of disorders associated with decreased LIPAP expression or activity, it is desirable to increase the expression or activity of LIPAP.

Therefore, in one embodiment, LIPAP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of LIPAP. Examples of such disorders include, but are not limited to, a cardiovascular disorder including blood vessel disorders such as arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebothrombosis, vascular tumors, and complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery; heart disorders such as congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, and complications of cardiac transplantation; and lung disorders such as congenital lung anomalies, atelectasis, pulmonary congestion and edema, pulmonary embolism, pulmonary hemorrhage, pulmonary infarction, pulmonary hypertension, vascular sclerosis, obstructive pulmonary disease, restrictive pulmonary disease, chronic obstructive pulmonary disease, emphysema, chronic bronchitis, bronchial asthma, bronchiectasis, bacterial pneumonia, viral and mycoplasmal pneumonia, lung abscess, pulmonary tuberculosis, diffuse interstitial diseases, pneumoconioses, sarcoidosis, idiopathic pulmonary fibrosis, desquamative interstitial pneumonitis, hypersensitivity pneumonitis, pulmonary eosinophilia bronchiolitis obliterans-organizing pneumonia, diffuse pulmonary hemorrhage syndromes, Goodpasture's syndromes, idiopathic pulmonary hemosiderosis, pulmonary involvement in collagen-vascular disorders, pulmonary alveolar proteinosis, lung tumors, inflammatory and noninflammatory pleural effusions, pneumothorax, pleural tumors, drug-induced lung disease, radiation-induced lung disease, and complications of lung transplantation; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders,

amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease; prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome; fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis; inherited, metabolic, endocrine, and toxic myopathies; myasthenia gravis, periodic paralysis; mental disorders including mood, anxiety, and schizophrenic disorders; seasonal affective disorder (SAD); akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; a gastrointestinal disorder such as dysphagia, peptic esophagitis, esophageal spasm, esophageal stricture, esophageal carcinoma, dyspepsia, indigestion, gastritis, gastric carcinoma, anorexia, nausea, emesis, gastroparesis, antral or pyloric edema, abdominal angina, pyrosis, gastroenteritis, intestinal obstruction, infections of the intestinal tract, peptic ulcer, cholelithiasis, cholecystitis, cholestasis, pancreatitis, pancreatic carcinoma, biliary tract disease, hepatitis, hyperbilirubinemia, cirrhosis, passive congestion of the liver, hepatoma, infectious colitis, ulcerative colitis, ulcerative proctitis, Crohn's disease, Whipple's disease, Mallory-Weiss syndrome, colonic carcinoma, colonic obstruction, irritable bowel syndrome, short bowel syndrome, diarrhea, constipation, gastrointestinal hemorrhage, acquired immunodeficiency syndrome (AIDS) enteropathy, jaundice, hepatic encephalopathy, hepatorenal syndrome, hepatic steatosis, hemochromatosis, Wilson's disease, alpha$_1$-antitrypsin deficiency, Reye's syndrome, primary sclerosing cholangitis, liver infarction, portal vein obstruction and thrombosis, centrilobular necrosis, peliosis hepatis, hepatic vein thrombosis, veno-occlusive disease, preeclampsia, eclampsia, acute fatty liver of pregnancy, intrahepatic cholestasis of pregnancy, and hepatic tumors including nodular hyperplasias, adenomas, and carcinomas; and a disorder of lipid metabolism such as fatty liver, cholestasis, primary biliary cirrhosis, carnitine deficiency, carnitine palmitoyltransferase deficiency, myoadenylate deaminase deficiency, hypertriglyceridemia, lipid storage disorders such Fabry's disease, Gaucher's disease, Niemann-Pick's disease, metachromatic leukodystrophy, adrenoleukodystrophy, GM_2 gangliosidosis, and ceroid lipofuscinosis, abetalipoproteinemia, Tangier disease, hyperlipoproteinemia, diabetes mellitus, lipodystrophy, lipomatoses, acute panniculitis, disseminated fat necrosis, adiposis dolorosa, lipoid adrenal hyperplasia, minimal change disease, lipomas,

atherosclerosis, hypercholesterolemia, hypercholesterolemia with hypertriglyceridemia, primary hypoalphalipoproteinemia, hypothyroidism, renal disease, liver disease, lecithin:cholesterol acyltransferase deficiency, cerebrotendinous xanthomatosis, sitosterolemia, hypocholesterolemia, Tay-Sachs disease, Sandhoff's disease, hyperlipidemia, hyperlipemia, lipid myopathies, and obesity.

In another embodiment, a vector capable of expressing LIPAP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of LIPAP including, but not limited to, those described above.

In a further embodiment, a pharmaceutical composition comprising a substantially purified LIPAP in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of LIPAP including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of LIPAP may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of LIPAP including, but not limited to, those listed above.

In a further embodiment, an antagonist of LIPAP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of LIPAP. Examples of such disorders include, but are not limited to, those cardiovascular, neurological, and gastrointestinal disorders, and disorders of lipid metabolism, described above. In one aspect, an antibody which specifically binds LIPAP may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express LIPAP.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding LIPAP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of LIPAP including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of LIPAP may be produced using methods which are generally known in the art. In particular, purified LIPAP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind LIPAP. Antibodies to LIPAP may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced

by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with LIPAP or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to LIPAP have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein and contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of LIPAP amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to LIPAP may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce LIPAP-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for LIPAP may also be generated. For example, such fragments include, but are not limited to, $F(ab')_2$ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between LIPAP and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering LIPAP epitopes is generally used, but a competitive binding assay may also be employed (Pound, supra).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for LIPAP. Affinity is expressed as an association constant, K_a, which is defined as the molar concentration of LIPAP-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple LIPAP epitopes, represents the average affinity, or avidity, of the antibodies for LIPAP. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular LIPAP epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the LIPAP-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of LIPAP, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington, DC; Liddell, J.E. and Cryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of LIPAP-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding LIPAP, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the polynucleotide encoding LIPAP may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding LIPAP. Thus, complementary molecules or fragments may be used to modulate LIPAP activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding LIPAP.

Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. Methods which are well known to those skilled in the art can be used to construct vectors to express nucleic acid sequences complementary to the polynucleotides encoding LIPAP. (See, e.g., Sambrook, supra; Ausubel, 1995, supra.)

Genes encoding LIPAP can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide, or fragment thereof, encoding LIPAP. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector, and may last even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5', or regulatory regions of the gene encoding LIPAP. Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may be employed. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze

endonucleolytic cleavage of sequences encoding LIPAP.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding LIPAP. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

An additional embodiment of the invention relates to the administration of a pharmaceutical or sterile composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of LIPAP,

antibodies to LIPAP, and mimetics, agonists, antagonists, or inhibitors of LIPAP. The compositions may be administered alone or in combination with at least one other agent, such as a stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs, or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA).

Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combining active compounds with solid excipient and processing the resultant mixture of granules (optionally, after grinding) to obtain tablets or dragee cores. Suitable auxiliaries can be added, if desired. Suitable excipients include carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, and sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums, including arabic and tragacanth; and proteins, such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, and alginic acid or a salt thereof, such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with fillers or binders, such as lactose or

starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils, such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate, triglycerides, or liposomes. Non-lipid polycationic amino polymers may also be used for delivery. Optionally, the suspension may also contain suitable stabilizers or agents to increase the solubility of the compounds and allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic acids. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preparation may be a lyophilized powder which may contain any or all of the following: 1 mM to 50 mM histidine, 0.1% to 2% sucrose, and 2% to 7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of LIPAP, such labeling would include amount, frequency, and method of administration.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example LIPAP or fragments thereof, antibodies of LIPAP, and agonists, antagonists or inhibitors of LIPAP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD_{50}/ED_{50} ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μg to 100,000 μg, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind LIPAP may be used for the diagnosis of disorders characterized by expression of LIPAP, or in assays to monitor patients being treated with LIPAP or agonists, antagonists, or inhibitors of LIPAP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for LIPAP include methods which utilize the antibody and a label to detect LIPAP in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring LIPAP, including ELISAs, RIAs, and FACS, are known in

the art and provide a basis for diagnosing altered or abnormal levels of LIPAP expression. Normal or standard values for LIPAP expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibody to LIPAP under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of LIPAP expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding LIPAP may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of LIPAP may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of LIPAP, and to monitor regulation of LIPAP levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding LIPAP or closely related molecules may be used to identify nucleic acid sequences which encode LIPAP. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding LIPAP, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the LIPAP encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:13-24 or from genomic sequences including promoters, enhancers, and introns of the LIPAP gene.

Means for producing specific hybridization probes for DNAs encoding LIPAP include the cloning of polynucleotide sequences encoding LIPAP or LIPAP derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding LIPAP may be used for the diagnosis of disorders associated with expression of LIPAP. Examples of such disorders include, but are not limited to, a cardiovascular disorder including blood vessel disorders such as arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins,

thrombophlebitis and phlebothrombosis, vascular tumors, and complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery; heart disorders such as congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, and complications of cardiac transplantation; and lung disorders such as congenital lung anomalies, atelectasis, pulmonary congestion and edema, pulmonary embolism, pulmonary hemorrhage, pulmonary infarction, pulmonary hypertension, vascular sclerosis, obstructive pulmonary disease, restrictive pulmonary disease, chronic obstructive pulmonary disease, emphysema, chronic bronchitis, bronchial asthma, bronchiectasis, bacterial pneumonia, viral and mycoplasmal pneumonia, lung abscess, pulmonary tuberculosis, diffuse interstitial diseases, pneumoconioses, sarcoidosis, idiopathic pulmonary fibrosis, desquamative interstitial pneumonitis, hypersensitivity pneumonitis, pulmonary eosinophilia bronchiolitis obliterans-organizing pneumonia, diffuse pulmonary hemorrhage syndromes, Goodpasture's syndromes, idiopathic pulmonary hemosiderosis, pulmonary involvement in collagen-vascular disorders, pulmonary alveolar proteinosis, lung tumors, inflammatory and noninflammatory pleural effusions, pneumothorax, pleural tumors, drug-induced lung disease, radiation-induced lung disease, and complications of lung transplantation; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease; prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome; fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis; inherited, metabolic, endocrine, and toxic myopathies; myasthenia gravis, periodic paralysis; mental disorders including mood, anxiety, and schizophrenic disorders; seasonal affective disorder (SAD); akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy,

corticobasal degeneration, and familial frontotemporal dementia; a gastrointestinal disorder such as dysphagia, peptic esophagitis, esophageal spasm, esophageal stricture, esophageal carcinoma, dyspepsia, indigestion, gastritis, gastric carcinoma, anorexia, nausea, emesis, gastroparesis, antral or pyloric edema, abdominal angina, pyrosis, gastroenteritis, intestinal obstruction, infections of the intestinal tract, peptic ulcer, cholelithiasis, cholecystitis, cholestasis, pancreatitis, pancreatic carcinoma, biliary tract disease, hepatitis, hyperbilirubinemia, cirrhosis, passive congestion of the liver, hepatoma, infectious colitis, ulcerative colitis, ulcerative proctitis, Crohn's disease, Whipple's disease, Mallory-Weiss syndrome, colonic carcinoma, colonic obstruction, irritable bowel syndrome, short bowel syndrome, diarrhea, constipation, gastrointestinal hemorrhage, acquired immunodeficiency syndrome (AIDS) enteropathy, jaundice, hepatic encephalopathy, hepatorenal syndrome, hepatic steatosis, hemochromatosis, Wilson's disease, alpha$_1$-antitrypsin deficiency, Reye's syndrome, primary sclerosing cholangitis, liver infarction, portal vein obstruction and thrombosis, centrilobular necrosis, peliosis hepatis, hepatic vein thrombosis, veno-occlusive disease, preeclampsia, eclampsia, acute fatty liver of pregnancy, intrahepatic cholestasis of pregnancy, and hepatic tumors including nodular hyperplasias, adenomas, and carcinomas; and a disorder of lipid metabolism such as fatty liver, cholestasis, primary biliary cirrhosis, carnitine deficiency, carnitine palmitoyltransferase deficiency, myoadenylate deaminase deficiency, hypertriglyceridemia, lipid storage disorders such Fabry's disease, Gaucher's disease, Niemann-Pick's disease, metachromatic leukodystrophy, adrenoleukodystrophy, GM_2 gangliosidosis, and ceroid lipofuscinosis, abetalipoproteinemia, Tangier disease, hyperlipoproteinemia, diabetes mellitus, lipodystrophy, lipomatoses, acute panniculitis, disseminated fat necrosis, adiposis dolorosa, lipoid adrenal hyperplasia, minimal change disease, lipomas, atherosclerosis, hypercholesterolemia, hypercholesterolemia with hypertriglyceridemia, primary hypoalphalipoproteinemia, hypothyroidism, renal disease, liver disease, lecithin:cholesterol acyltransferase deficiency, cerebrotendinous xanthomatosis, sitosterolemia, hypocholesterolemia, Tay-Sachs disease, Sandhoff's disease, hyperlipidemia, hyperlipemia, lipid myopathies, and obesity. The polynucleotide sequences encoding LIPAP may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered LIPAP expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding LIPAP may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding LIPAP may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control

sample then the presence of altered levels of nucleotide sequences encoding LIPAP in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of LIPAP, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding LIPAP, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding LIPAP may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding LIPAP, or a fragment of a polynucleotide complementary to the polynucleotide encoding LIPAP, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

Methods which may also be used to quantify the expression of LIPAP include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236.) The speed of quantitation of multiple samples may be

accelerated by running the assay in a high-throughput format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, and to develop and monitor the activities of therapeutic agents.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.)

In another embodiment of the invention, nucleic acid sequences encoding LIPAP may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical chromosome mapping techniques and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding LIPAP on a physical chromosomal map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder. The nucleotide sequences of the invention may be used to detect differences in gene sequences among normal, carrier, and affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene

discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, LIPAP, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between LIPAP and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with LIPAP, or fragments thereof, and washed. Bound LIPAP is then detected by methods well known in the art. Purified LIPAP can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding LIPAP specifically compete with a test compound for binding LIPAP. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with LIPAP.

In additional embodiments, the nucleotide sequences which encode LIPAP may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, in particular U.S. Ser. No. 60/120,703 and U.S. Ser. No. 60/142,762, are hereby expressly incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

RNA was purchased from Clontech or isolated from tissues described in Table 4. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A+) RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), or pINCY (Incyte Pharmaceuticals, Palo Alto CA). Recombinant plasmids were transformed into competent E. coli cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5α, DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids were recovered from host cells by in vivo excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or

without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

cDNA sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Perkin-Elmer) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Perkin-Elmer) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, supra, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VI.

The polynucleotide sequences derived from cDNA sequencing were assembled and analyzed using a combination of software programs which utilize algorithms well known to those skilled in the art. Table 5 summarizes the tools, programs, and algorithms used and provides applicable descriptions, references, and threshold parameters. The first column of Table 5 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score, the greater the homology between two sequences). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments were generated using the default parameters specified by the clustal algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

The polynucleotide sequences were validated by removing vector, linker, and polyA sequences

and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programing, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and PFAM to acquire annotation using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and Hidden Markov Model (HMM)-based protein family databases such as PFAM. HMM is a probabilistic approach which analyzes consensus primary structures of gene families. (See, e.g., Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.)

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:13-24. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

IV. Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; Ausubel, 1995, supra, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in nucleotide databases such as GenBank or LIFESEQ (Incyte Pharmaceuticals). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\%\ \text{sequence identity} \times \%\ \text{maximum BLAST score}}{100}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1% to 2% error, and, with a product score of 70, the match will be exact. Similar molecules are usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

The results of northern analyses are reported as a percentage distribution of libraries in which

the transcript encoding LIPAP occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, reproductive, and urologic. The disease/condition categories included cancer, inflammation, trauma, cell proliferation, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease- or condition-specific expression are reported in Table 3.

V. Chromosomal Mapping of LIPAP Encoding Polynucleotides

The cDNA sequences which were used to assemble SEQ ID NO:19-24 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:19-24 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 5). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO:, to that map location.

The genetic map location of SEQ ID NO:21 is described in The Invention as a range, or interval, of human chromosome 11. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Généthon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters. Human genome maps and other resources available to the public, such as the NCBI "GeneMap'99" World Wide Web site (http://www.ncbi.nlm.nih.gov/genemap/), can be employed to determine if previously identified disease genes map within or in proximity to the intervals indicated above.

VI. Extension of LIPAP Encoding Polynucleotides

The full length nucleic acid sequences of SEQ ID NO:13-24 were produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at

temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+}, $(NH_4)_2SO_4$, and β-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 µl PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 µl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 μl to 10 μl aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent E. coli cells. Transformed cells were selected on antibiotic-containing media, individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by

PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethysulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer).

In like manner, the nucleotide sequences of SEQ ID NO:13-24 are used to obtain 5' regulatory sequences using the procedure above, oligonucleotides designed for such extension, and an appropriate genomic library.

VII. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:13-24 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μCi of [γ-^{32}P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

VIII. Microarrays

A chemical coupling procedure and an ink jet device can be used to synthesize array elements on the surface of a substrate. (See, e.g., Baldeschweiler, supra.) An array analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced by hand or using available methods and machines and contain any appropriate number of elements. After hybridization, nonhybridized probes are removed and a scanner used to determine the levels and patterns of fluorescence. The degree of complementarity and the relative abundance of each probe which hybridizes to an element on the microarray may be assessed through analysis of the scanned images.

Full-length cDNAs, Expressed Sequence Tags (ESTs), or fragments thereof may comprise the elements of the microarray. Fragments suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). Full-length cDNAs, ESTs, or fragments thereof corresponding to one of the nucleotide sequences of the present invention, or selected at random from a cDNA library relevant to the present invention, are arranged on an appropriate substrate, e.g., a glass slide. The cDNA is fixed to the slide using, e.g., UV cross-linking followed by thermal and chemical treatments and subsequent drying. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645.) Fluorescent probes are prepared and used for hybridization to the elements on the substrate. The substrate is analyzed by procedures described above.

IX. Complementary Polynucleotides

Sequences complementary to the LIPAP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring LIPAP. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of LIPAP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the LIPAP-encoding transcript.

X. Expression of LIPAP

Expression and purification of LIPAP is achieved using bacterial or virus-based expression systems. For expression of LIPAP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express LIPAP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of LIPAP in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding LIPAP by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et

al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, LIPAP is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from LIPAP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch. 10 and 16). Purified LIPAP obtained by these methods can be used directly in the following activity assay.

XI. Demonstration of LIPAP Activity

Selected candidate lipid molecules, such as C4 sterols, oxysterol, apolipoprotein E, and phospholipids, are arrayed in the wells of a multi-well plate. LIPAP, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) Biochem. J. 133:529-539.) The selected candidate lipid molecules are incubated with the labeled LIPAP and washed. Any wells with labeled LIPAP complex are assayed. Data obtained using different concentrations of LIPAP are used to calculate values for the number, affinity, and association of LIPAP with the candidate molecules. Significant binding of LIPAP to the candidate lipid molecules is indicative of LIPAP activity.

XII. Functional Assays

LIPAP function is assessed by expressing the sequences encoding LIPAP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT (Life Technologies) and pCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μg of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μg of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-

based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of LIPAP on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding LIPAP and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding LIPAP and other genes of interest can be analyzed by northern analysis or microarray techniques.

XIII. Production of LIPAP Specific Antibodies

LIPAP substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) Methods Enzymol. 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the LIPAP amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Perkin-Elmer) using fmoc-chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-LIPAP activity by, for example, binding the peptide or LIPAP to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XIV. Purification of Naturally Occurring LIPAP Using Specific Antibodies

Naturally occurring or recombinant LIPAP is substantially purified by immunoaffinity chromatography using antibodies specific for LIPAP. An immunoaffinity column is constructed by covalently coupling anti-LIPAP antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing LIPAP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of LIPAP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/LIPAP binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and LIPAP is collected.

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
1	13	161190	ADENINB01	161190H1 (ADENINB01), 161190R6 (ADENINB01), 686052H1 (UTRSNOT02), 1962050R6 (BRSTNOT04), 2318534H1 (OVARNOT02), 2583728F6 (BRAITUT22), 3043537H1 (HEAANOT01), 3684806T6 (HEAANOT01), 4047892H1 (LUNGNOT35), 4335403F6 (KIDCTMT01)
2	14	1292575	PGANNOT03	982431T2 (TONGTUT01), 996331R1 (KIDNTUT01), 1292575H1 (PGANNOT03), 1478462F1 (CORPNOT02), 1478462T1 (CORPNOT02), 1731035F6 (BRSTTUT08), 1752672F6 (LIVRTUT01), 1752672T6 (LIVRTUT01), 2046050F6 (THP1T7T01), 3111288H1 (BRSTNOT17), 5293851H1 (COLENOT01)
3	15	2454393	ENDANOT01	548115F1 (BEPINOT01), 2454393H1 (ENDANOT01), 3176463T6 (UTRSTUT04), 3742952H1 (THYMNOT08), 4415344H1 (MONOTXT01), SBIA04456D1, SBIA00571D1, SBIA03488D1, SBIA02429D1
4	16	2766980	BRSTNOT12	027244F1 (SPLNFET01), 084571H1 (HYPONOB01), 150574F1 (FIBRANT01), 237612R1 (SINTNOT02), 269891X13 (HNT2NOT01), 416256R1 (BRSTNOT01), 1345369F6 (PROSNOT11), 2766980H1 (BRSTNOT12), 2766980X305D1 (BRSTNOT12), 2806266H1 (BLADTUT08), 4200618H1 (BRAITUT29)
5	17	2768356	COLANOT02	1381442F6 (BRAITUT08), 2120949T6 (BRSTNOT07), 2768356H1 (COLANOT02), 2796651F6 (NPOLNOT01)
6	18	5324145	FIBPFEN06	638172F1 (BRSTNOT03), 1440822F6 (THYRNOT03), 1559428F6 (SPLNNOT04), 2236370F6 (PANCTUT02), 5324145H1 (FIBPFEN06)
7	19	1004646	BRSTNOT03	1004646H1 (BRSTNOT03), 1004646X312D1 (BRSTNOT03), SCHA04882V1, SBHA00389F1, SCIA00646V1
8	20	1802851	COLNNOT27	1811146F6 (PROSTUT12), 2347065T6 (TESTTUT02), 2500280F6 (ADRETUT05), 3076454H1 (BONEUNT01)

Table 1 (cont.)

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
9	21	2764333	BRSTNOT12	661437R6 (BRAINOT03), 1551790R6 (PROSNOT06), 1800152T6 (COLNNOT27), 2123941F6 (BRSTNOT07), 2123941T6 (BRSTNOT07), 2185882H1 (PROSNOT26), 2764333H1 (BRSTNOT12), 2764333T6 (BRSTNOT12), 4616050H1 (BRAYDIT01)
10	22	2798021	NPOLNOT01	1311367F1 (COLNFET02), 1458887T6.com (COLNFET02), 2798021F6 (NPOLNOT01), 2798021H1 (NPOLNOT01), 2798021T6.com (NPOLNOT01), 2936035F6 (THYMFET02), 2936035T6.com (THYMFET02)
11	23	3335404	BRAIFET01	090725H1 (HYPONOB01), 1440011F1 (THYRNOT03), 1593543F1 (BRAINOT14), 1593543T6 (BRAINOT14), 2552343T6 (LUNGTUT06), 2783819H1 (BRSTNOT13), 2885772F6 (SINJNOT02), 2885772T6 (SINJNOT02), 3335404H1 (BRAIFET01)
12	24	3735780	SMCCNOS01	551126H1 (BEPINOT01), 2808373H1 (BLADTUT08), 3735780F6 (SMCCNOS01), 3735780H1 (SMCCNOS01), 3735780T6 (SMCCNOS01), 4760604T6 (BRAMNOT01)

Table 2

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences	Identification	Analytical Methods and Databases
1	331	T9 T30 T40 S167 S175 T187 S300 S3 T240 T258 Y15 Y53			Apolipoprotein L (g2425058)	BLAST
2	480	T12 S19 S70 S143 S385 T391 T38 T148 S316 S336	N18 N378 N472	Oxysterol-binding protein: D35-N472	Oxysterol-binding protein (g3551523)	MOTIFS PFAM BLOCKS BLAST
3	409	S188 S268 S348 T358 S238 S275 S328 S341	N164	Phosphatidylserine decarboxylase: H161-K174, Y257-P269, N325-H338, G368-F384	Phosphatidylserine decarboxylase (g191185)	BLIMPS-PRODOM BLAST
4	759	S55 T68 S225 S582 T19 T48 S85 S93 S132 S168 S230 S244 S266 S294 T318 S326 T337 S369 T389 S467 S514 S543 S563 S583 S617 S658 S686 S698 S709 T714 S741 S15 S89 S158 S184 S220 S248 S253 T525 S601 S604 S642 T662 Y229	N29 N59 N92 N251 N286 N706	LIM domain: R344-Q444	Sterol regulatory element binding protein-2 (g841318)	MOTIFS ProfileScan PFAM BLOCKS BLAST
5	226	T77 S197 T207 S218 S82 S137 Y56 Y98 Y122	N205	Oxysterol-binding protein: D146-H189	Oxysterol-binding protein	BLOCKS
6	500	S402 T14 S50 S52 T80 S242 T254 T403 T473 S46 T106 S244 S435 Y266	N78 N104 N433		Niemann-Pick C disease-associated gene product (g2251248)	BLAST

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences	Identification	Analytical Methods and Databases
7	272	S175 S211 T255 T224 T263	N5 N163 N189	Transmembrane sterol biosynthesis oxidoreductase: Y48-T263	C4 sterol methyl oxidase (g1161339)	MOTIFS BLAST-GenBank BLIMPS-PRODOM
8	282	S84 T140 S161 S218 T71 T95 T120 T149 S192	N126 N195 N213	Signal peptide: M1-G30 Prokaryotic membrane lipoprotein attachment site: A44-C54 Leucine zipper motif: L17-L38 Low-density lipoprotein receptor domain class A: G52-I91, C152-E164	Apolipoprotein E receptor 2 (g1834534)	MOTIFS BLAST-GenBank HMMER-PFAM SigPept BLIMPS-BLOCKS BLAST-PRODOM
9	437	T285 S291 S30 S82 S103 S295 T296 S395 S397 S424 T108 S251 Y267	N335 N393	Oxysterol-binding protein signature: E134-A144, K18-P271	Oxysterol-binding protein (g189403)	MOTIFS BLAST-GenBank HMMER-PFAM BLIMPS-BLOCKS BLAST-DOMO
10	427	T333 T199 T212 S224 T233 S281 S295 T29 S103 S128 S249 S292 Y151 Y339	N132 N293	ATP-binding protein motif: D146-K372	CDV-1R protein suppressed in steatosis (g2995447)	MOTIFS BLAST-GenBank BLAST-PRODOM
11	564	T280 T24 S42 S108 T119 S129 T149 S187 S251 S297 S428 S31 S216 T228 S356 S403 S490	N6 N106 N309 N458	C2 (protein kinase C) domain: A166-F183 Synaptotagmin protein kinase motif: A166-S282	Copine I (phospholipid-binding protein) (g1791257)	MOTIFS BLAST-GenBank HMMER-PFAM BLAST-PRODOM

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences	Identification	Analytical Methods and Databases
12	297	S17 S114 T136 S16	N287	Mitochondrial carrier protein: E117-I297 Graves disease carrier protein motif: P137-T157	Similar to human ADP/ATP carrier protein (g3879938)	MOTIFS BLAST-GenBank HMMER-PFAM BLIMPS-PRINTS

Table 3

SEQ ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
13	433-477	Reproductive (0.250) Gastrointestinal (0.183) Cardiovascular (0.144)	Cell Proliferation (0.500) Inflammation (0.500)	PBLUESCRIPT
14	202-270	Reproductive (0.250) Nervous (0.219) Gastrointestinal (0.141)	Cell Proliferation (0.531) Inflammation (0.313)	pINCY
15	731-802	Reproductive (0.235) Nervous (0.176) Gastrointestinal (0.153) Cardiovascular (0.129)	Cell Proliferation (0.553) Inflammation (0.388)	PBLUESCRIPT
16	875-919 1544-1609	Reproductive (0.296) Gastrointestinal (0.178) Nervous (0.118)	Cell Proliferation (0.632) Inflammation (0.276)	pINCY
17	96-155	Gastrointestinal (0.214) Hematopoietic/Immune (0.214) Reproductive (0.214) Cardiovascular (0.143) Nervous (0.143)	Cell Proliferation (0.643) Inflammation (0.357)	pINCY
18	1075-1257	Hematopoietic/Immune (0.256) Gastrointestinal (0.179) Reproductive (0.154) Cardiovascular (0.128)	Cell Proliferation (0.513) Inflammation (0.436)	pINCY

Table 3 (cont.)

SEQ ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
19	281-325	Reproductive (0.250) Gastrointestinal (0.250) Hematopoietic/Immune (0.125) Nervous (0.125)	Cancer (0.313) Inflammation/Trauma (0.376) Cell proliferation (0.124)	PSPORT1
20	218-262	Reproductive (0.270) Nervous (0.175) Gastrointestinal (0.095)	Cancer (0.508) Cell proliferation (0.238) Inflammation/Trauma (0.238)	pINCY
21	279-326	Nervous (0.302) Reproductive (0.281) Gastrointestinal (0.146)	Cancer (0.427) Inflammation/Trauma (0.323) Cell proliferation (0.094)	pINCY
22	55-99	Reproductive (0.286) Nervous (0.200) Cardiovascular (0.114) Developmental (0.114)	Cell proliferation (0.286) Cancer (0.286) Inflammation/Trauma (0.246)	pINCY
23	434-478	Nervous (0.318) Cardiovascular (0.182) Gastrointestinal (0.136)	Cancer (0.500) Cell proliferation (0.227) Inflammation/Trauma (0.272)	pINCY
24	219-263	Nervous (0.231) Cardiovascular (0.231) Hematopoietic/Immune (0.154) Reproductive (0.154)	Cancer (0.231) Cell proliferation (0.154) Inflammation/Trauma (0.154)	pINCY

Table 4

Polynucleotide SEQ ID NO:	Library	Library Comment
13	ADENINB01	Library was constructed using RNA isolated from the inflamed adenoid tissue of a 3-year-old child. The RNA came from Clontech.
14	PGANNOT03	Library was constructed using RNA isolated from paraganglionic tumor tissue removed from the intra-abdominal region of a 46-year-old Caucasian male during exploratory laparotomy. Pathology indicated a benign paraganglioma and was associated with a grade 2 renal cell carcinoma, clear cell type, which did not penetrate the capsule.
15	ENDANOT01	Library was constructed using RNA isolated from aortic endothelial cell tissue from an explanted heart removed from a male during a heart transplant.
16	BRSTNOT12	Library was constructed using RNA isolated from diseased breast tissue removed from a 32-year-old Caucasian female during a bilateral reduction mammoplasty. Pathology indicated nonproliferative fibrocystic disease. Family history included benign hypertension and atherosclerotic coronary artery disease.
17	COLANOT02	Library was constructed using RNA isolated from diseased ascending colon tissue removed from a 25-year-old Caucasian female during a multiple segmental resection of the large bowel. Pathology indicated moderately to severely active chronic ulcerative colitis, involving the entire colectomy specimen and sparing 2 cm of the attached ileum. Grossly, the specimen showed continuous involvement from the rectum proximally; marked mucosal atrophy and no skip areas were identified. Microscopically, the specimen showed dense, predominantly mucosal inflammation and crypt abscesses. Patient history included benign large bowel neoplasm.
18	FIBPFEN06	This normalized prostate stromal fibroblast tissue library was constructed from 1.56 million independent clones from a fibroblast library. Starting RNA was made from fibroblasts of prostate stroma removed from a male fetus, who died after 26 weeks' gestation. The libraries were normalized in two rounds using conditions adapted from Soares et al. (1994) Proc. Natl. Acad. Sci. USA 91:9228 and Bonaldo et al.(1996) Genome Research 6:791, except that a significantly longer (48-hours/round) reannealing hybridization was used.

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Library	Library Comment
19	BRSTNOT03	Library was constructed using RNA isolated from diseased breast tissue removed from a 54-year-old Caucasian female during a bilateral radical mastectomy. Pathology for the associated tumor tissue indicated residual invasive grade 3 mammary ductal adenocarcinoma. Patient history included kidney infection and condyloma acuminatum. Family history included benign hypertension, hyperlipidemia and a malignant neoplasm of the colon.
20	COLNNOT27	Library was constructed using RNA isolated from diseased cecal tissue removed from 31-year-old Caucasian male during a total intra-abdominal colectomy, appendectomy, and permanent ileostomy. Pathology indicated severe active Crohn's disease involving the colon from the cecum to the rectum.
21	BRSTNOT12	Library was constructed using RNA isolated from diseased breast tissue removed from a 32-year-old Caucasian female during a bilateral reduction mammoplasty. Pathology indicated nonproliferative fibrocystic disease. Family history included benign hypertension and atherosclerotic coronary artery disease.
22	NPOLNOT01	Library was constructed using RNA isolated from nasal polyp tissue removed from a 78-year-old Caucasian male during a nasal polypectomy. Pathology indicated a nasal polyp and striking eosinophilia. Patient history included asthma and nasal polyps.
23	BRAIFET01	Library was constructed using RNA isolated from brain tissue removed from a Caucasian male fetus, who was stillborn with a hypoplastic left heart at 23 weeks' gestation.
24	SMCCNOS01	This subtracted coronary artery smooth muscle cell library was constructed using 7.56 x 10e6 clones from a treated coronary artery smooth muscle cell library and was subjected to two rounds of subtraction hybridization for 48 hours with 6.12 x 10e6 clones from an untreated coronary artery smooth muscle cell library. The starting library for subtraction was constructed using RNA isolated from coronary artery smooth muscle cells removed from a 3-year-old Caucasian male. The cells were treated with TNF alpha and IL-1 beta, 10ng/ml each, for 20 hours. The hybridization probe for subtraction was derived from a similarly constructed library from RNA isolated from untreated coronary artery smooth muscle cells from the same donor. Subtractive hybridization conditions were based on the methodologies of Swaroop et al. (1991) Nucleic Acids Res. 19:1954 and Bonaldo et al. (1996) Genome Research 6:791-806.

Table 5

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25: 3389-3402.	*ESTs:* Probability value= 1.0E-8 or less *Full Length sequences*: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises as least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; and Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	*ESTs:* fasta E value=1.06E-6 *Assembled ESTs*: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less *Full Length sequences:* fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S and J.G. Henikoff, Nucl. Acid Res., 19:6565-72, 1991. J.G. Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and, if applicable, Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322.	Score=10-50 bits for PFAM hits, depending on individual protein families

Table 5 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A et al. (1997) Nucleic Acids Res. 25: 217-221.	Normalized quality score≥GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	Score=3.5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch et al. supra; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

a) an amino acid sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, and SEQ ID NO:12,

b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, and SEQ ID NO:12,

c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, and SEQ ID NO:12, and

d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, and SEQ ID NO:12.

2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, and SEQ ID NO:12.

3. An isolated polynucleotide encoding a polypeptide of claim 1.

4. An isolated polynucleotide of claim 3 selected from the group consisting of SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, and SEQ ID NO:24.

5. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.

6. A cell transformed with a recombinant polynucleotide of claim 5.

7. A transgenic organism comprising a recombinant polynucleotide of claim 5.

8. A method for producing a polypeptide of claim 1, the method comprising:

a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and

b) recovering the polypeptide so expressed.

9. An isolated antibody which specifically binds to a polypeptide of claim 1.

10. An isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of:

a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, and SEQ ID NO:24,

b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, and SEQ ID NO:24,

c) a polynucleotide sequence complementary to a),

d) a polynucleotide sequence complementary to b), and

e) an RNA equivalent of a)-d).

11. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 10.

12. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 10, the method comprising:

a) hybridizing the sample with a probe comprising at least 16 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide, and

b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

13. A method of claim 12, wherein the probe comprises at least 30 contiguous nucleotides.

14. A method of claim 12, wherein the probe comprises at least 60 contiguous nucleotides.

15. A pharmaceutical composition comprising an effective amount of a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

16. A method for treating a disease or condition associated with decreased expression of functional LIPAP, comprising administering to a patient in need of such treatment the pharmaceutical composition of claim 15.

17. A method for screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting agonist activity in the sample.

18. A pharmaceutical composition comprising an agonist compound identified by a method of claim 17 and a pharmaceutically acceptable excipient.

19. A method for treating a disease or condition associated with decreased expression of functional LIPAP, comprising administering to a patient in need of such treatment a pharmaceutical composition of claim 18.

20. A method for screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting antagonist activity in the sample.

21. A pharmaceutical composition comprising an antagonist compound identified by a method of claim 20 and a pharmaceutically acceptable excipient.

22. A method for treating a disease or condition associated with overexpression of functional LIPAP, comprising administering to a patient in need of such treatment a pharmaceutical composition of claim 21.

23. A method for screening a compound for effectiveness in altering expression of a target

polynucleotide, wherein said target polynucleotide comprises a sequence of claim 4, the method comprising:

a) exposing a sample comprising the target polynucleotide to a compound, and

b) detecting altered expression of the target polynucleotide.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C07K 14/00 — A2

(11) International Publication Number: WO 00/49043

(43) International Publication Date: 24 August 2000 (24.08.00)

(21) International Application Number: PCT/US00/04160

(22) International Filing Date: 18 February 2000 (18.02.00)

(30) Priority Data:

60/120,703	19 February 1999 (19.02.99)	US
60/142,762	8 July 1999 (08.07.99)	US

(71) Applicant *(for all designated States except US)*: INCYTE PHARMACEUTICALS, INC. [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: TANG, Y. Tom [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). HILLMAN, Jennifer, L. [US/US]; 230 Monroe Drive #12, Mountain View, CA 94040 (US). YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). AZIMZAI, Yalda [US/US]; 2045 Rock Springs Drive, Hayward, CA 94545 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, Sunnyvale, CA 94577 (US). TRAN, Bao [US/US]; 744 Kiely Boulevard, Santa Clara, CA 95051 (US).

(74) Agents: HAMLET–COX, Diana et al.; Incyte Pharmaceuticals, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: HUMAN LIPID–ASSOCIATED PROTEINS

(57) Abstract

The invention provides human lipid–associated proteins (LIPAP) and polynucleotides which identify and encode LIPAP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of LIPAP.

Docket No.: PF-0676 USN

DECLARATION AND POWER OF ATTORNEY FOR UNITED STATES PATENT APPLICATION

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name, and

I believe that I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if more than one name is listed below) of the subject matter which is claimed and for which a United States patent is sought on the invention entitled

HUMAN LIPID-ASSOCIATED PROTEINS

the specification of which:

/___/ is attached hereto.

/___/ was filed on _____ as application Serial No. _____ and if this box contains an X /___/, was amended on _____.

/ X / was filed as Patent Cooperation Treaty international application No. PCT/US00/04160 on February 18, 2000, if this box contains an X /_/, was amended on under Patent Cooperation Treaty Article 19 on _____ 2001, and if this box contains an X /_/, was amended on _____.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge my duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, §1.56(a).

I hereby claim the benefit under Title 35, United States Code, §119 or §365(a)-(b) of any foreign application(s) for patent or inventor's certificate indicated below and of any Patent Cooperation Treaty international applications(s) designating at least one country other than the United States indicated below and have also identified below any foreign application(s) for patent or inventor's certificate and Patent Cooperation Treaty international application(s) designating at least one country other than the United States for the same subject matter and having a filing date before that of the application for said subject matter the priority of which is claimed:

Docket No.: PF-0676 USN

Country	Number	Filing Date	Priority Claimed
_____	_____	_____	/_/ Yes /_/ No
_____	_____	_____	/_/ Yes /_/ No

I hereby claim the benefit under Title 35, United States Code, §119(e) of any United States provisional application(s) listed below.

Application Serial No.	Filed	Status (Pending, Abandoned, Patented)
60/120,703	February 19, 1999	Expired
60/142,762	July 8, 1999	Expired

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in said prior application(s) in the manner required by the first paragraph of Title 35, United States Code §112, I acknowledge my duty to disclose material information as defined in Title 37 Code of Federal Regulations, §1.56(a) which occurred between the filing date(s) of the prior application(s) and the national or Patent Cooperation Treaty international filing date of this application:

Application Serial No.	Filed	Status (Pending, Abandoned, Patented)

I hereby appoint the following:

Lucy J. Billings	Reg. No. 36,749
Michael C. Cerrone	Reg. No. 39,132
Diana Hamlet-Cox	Reg. No. ~~33,302~~
Richard C. Ekstrom	Reg. No. 37,027
Barrie D. Greene	Reg. No. 46,740
Lynn E. Murry	Reg. No. ~~42,918~~
Shirley A. Recipon	Reg. No. ~~47,016~~
Susan K. Sather	Reg. No. 44,316
Michelle M. Stempien	Reg. No. ~~41,327~~
David G. Streeter	Reg. No. 43,168
P. Ben Wang	Reg. No. 41,420

respectively and individually, as my patent attorneys and/or agents, with full power of substitution and revocation, to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith. Please address all communications to:

Docket No.: PF-0676 USN

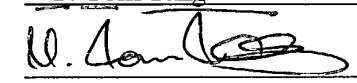
LEGAL DEPARTMENT
INCYTE GENOMICS, INC.
3160 PORTER DRIVE, PALO ALTO, CA 94304

TEL: 650-855-0555 FAX: 650-849-8886 or 650-845-4166

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

1-00 **First Joint Inventor:**

Full name: Y. Tom Tang

Signature: 

Date: September 10, 2001

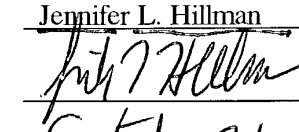
Citizenship: United States

Residence: San Jose, California CA

P.O. Address: 4230 Ranwick Court
San Jose, California 95118

3-00 **Second Joint Inventor:**

Full name: Jennifer L. Hillman

Signature: 

Date: September 21, 2001

Citizenship: United States

Residence: Mountain View, California CA

P.O. Address: 230 Monroe Drive, #17
Mountain View, California 94040

3-00 **Third Joint Inventor:**

Full name: Henry Yue

Signature: Henry Yue.

Date: September 24, 2001

Citizenship: United States

Residence: Sunnyvale, California CA

P.O. Address: 826 Lois Avenue
Sunnyvale, California 94087

4-00 **Fourth Joint Inventor:**

Full name: Yalda Azimzai

Signature: Yalda Azimzai

Date: September 13, 2001

Citizenship: United States

Residence: Castro Valley, California CA

P.O. Address: 5518 Boulder Canyon Drive
Castro Valley, California 94552

5-00 **Fifth Joint Inventor:**

Full name: Mariah R. Baughn

Signature: Mariah R. Baughn

Date: September 5, 2001

Citizenship: United States

Residence: San Leandro, California CA

P.O. Address: 14244 Santiago Road
San Leandro, California 94577

6-a0 **Sixth Joint Inventor:**

Full name: Bao Tran

Signature: _____

Date: _____, 2001

Citizenship: United States

Residence: Santa Clara, California CA

P.O. Address: 750 Salberg Avenue
San Leandro, California 95051

Docket No.: PF-0676 USN

DECLARATION AND POWER OF ATTORNEY FOR UNITED STATES PATENT APPLICATION

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name, and

I believe that I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if more than one name is listed below) of the subject matter which is claimed and for which a United States patent is sought on the invention entitled

HUMAN LIPID-ASSOCIATED PROTEINS

the specification of which:

/___/ is attached hereto.

/___/ was filed on _____ as application Serial No. _____ and if this box contains an X /___/, was amended on _____.

/ X / was filed as Patent Cooperation Treaty international application No. PCT/US00/04160 on February 18, 2000, if this box contains an X /_/, was amended on under Patent Cooperation Treaty Article 19 on _____ 2001, and if this box contains an X /_/, was amended on _____.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge my duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, §1.56(a).

I hereby claim the benefit under Title 35, United States Code, §119 or §365(a)-(b) of any foreign application(s) for patent or inventor's certificate indicated below and of any Patent Cooperation Treaty international applications(s) designating at least one country other than the United States indicated below and have also identified below any foreign application(s) for patent or inventor's certificate and Patent Cooperation Treaty international application(s) designating at least one country other than the United States for the same subject matter and having a filing date before that of the application for said subject matter the priority of which is claimed:

Docket No.: PF-0676 USN

Country	Number	Filing Date	Priority Claimed
_____	_____	_____	/_/ Yes /_/ No
_____	_____	_____	/_/ Yes /_/ No

I hereby claim the benefit under Title 35, United States Code, §119(e) of any United States provisional application(s) listed below.

Application Serial No.	Filed	Status (Pending, Abandoned, Patented)
60/120,703	February 19, 1999	Expired
60/142,762	July 8, 1999	Expired

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in said prior application(s) in the manner required by the first paragraph of Title 35, United States Code §112, I acknowledge my duty to disclose material information as defined in Title 37 Code of Federal Regulations, §1.56(a) which occurred between the filing date(s) of the prior application(s) and the national or Patent Cooperation Treaty international filing date of this application:

Application Serial No.	Filed	Status (Pending, Abandoned, Patented)

I hereby appoint the following:

Lucy J. Billings	Reg. No. 36,749
Michael C. Cerrone	Reg. No. 39,132
Diana Hamlet-Cox	Reg. No. 33,302
Richard C. Ekstrom	Reg. No. 37,027
Barrie D. Greene	Reg. No. 46,740
Lynn E. Murry	Reg. No. 42,918
Shirley A. Recipon	Reg. No. 47,016
Cathleen M. Rocco	Reg. No. 46,172
Susan K. Sather	Reg. No. 44,316
Michelle M. Stempien	Reg. No. 41,327
David G. Streeter	Reg. No. 43,168

respectively and individually, as my patent attorneys and/or agents, with full power of substitution and revocation, to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith. Please address all communications to:

LEGAL DEPARTMENT
INCYTE GENOMICS, INC.
3160 PORTER DRIVE, PALO ALTO, CA 94304

TEL: 650-855-0555 FAX: 650-849-8886 or 650-845-4166

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

First Joint Inventor:

Full name: Y. Tom Tang

Signature: _____

Date: _____, 2001

Citizenship: United States

Residence: San Jose, California

P.O. Address: 4230 Ranwick Court
San Jose, California 95118

Second Joint Inventor:

Full name: Jennifer L. Hillman

Signature: _____

Date: _____, 2001

Citizenship: United States

Residence: Mountain View, California

P.O. Address: 230 Monroe Drive, #17
Mountain View, California 94040

Third Joint Inventor:

Full name: Henry Yue

Signature: _____

Date: _____, 2001

Citizenship: United States

Residence: Sunnyvale, California

P.O. Address: 826 Lois Avenue
Sunnyvale, California 94087

Fourth Joint Inventor:

Full name: Yalda Azimzai

Signature: _____

Date: _____, 2001

Citizenship: United States

Residence: Castro Valley, California

P.O. Address: 5518 Boulder Canyon Drive
Castro Valley, California 94552

Fifth Joint Inventor:

Full name: Mariah R. Baughn

Signature: _____

Date: _____, 2001

Citizenship: United States

Residence: San Leandro, California

P.O. Address: 14244 Santiago Road
San Leandro, California 94577

Sixth Joint Inventor:

Full name: Bao Tran

Signature:

Date: Oct 24, 2001

Citizenship: United States

Residence: Santa Clara, California

P.O. Address: 750 Salberg Avenue
San Leandro, California 95051

SEQUENCE LISTING

<110> INCYTE PHARMACEUTICALS, INC.
TANG, Y. Tom
HILLMAN, Jennifer L.
YUE, Henry
AZIMZAI, Yalda
BAUGHN, Mariah R.
TRAN, Bao

<120> HUMAN LIPID-ASSOCIATED PROTEINS

<130> PF-0676 PCT

<140> To Be Assigned
<141> Herewith

<150> 60/120,703; 60/142,762
<151> 1999-02-19; 1999-07-08

<160> 24

<170> PERL Program

<210> 1
<211> 331
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 161190CD1

<400> 1

Met	Asp	Ser	Glu	Lys	Lys	Arg	Phe	Thr	Glu	Glu	Ala	Thr	Lys	Tyr	
1				5				10						15	
Phe	Arg	Glu	Arg	Val	Ser	Pro	Val	His	Leu	Gln	Ile	Leu	Leu	Thr	
				20				25						30	
Asn	Asn	Glu	Ala	Trp	Lys	Arg	Phe	Val	Thr	Ala	Ala	Glu	Leu	Pro	
				35				40						45	
Arg	Asp	Glu	Ala	Asp	Ala	Leu	Tyr	Glu	Ala	Leu	Lys	Lys	Leu	Arg	
				50				55						60	
Thr	Tyr	Ala	Ala	Ile	Glu	Asp	Glu	Tyr	Val	Gln	Gln	Lys	Asp	Glu	
				65				70						75	
Gln	Phe	Arg	Glu	Trp	Phe	Leu	Lys	Glu	Phe	Pro	Gln	Val	Lys	Arg	
				80				85						90	
Lys	Ile	Gln	Glu	Ser	Ile	Glu	Lys	Leu	Arg	Ala	Leu	Ala	Asn	Gly	
				95				100						105	
Ile	Glu	Glu	Val	His	Arg	Gly	Cys	Thr	Ile	Ser	Asn	Val	Val	Ser	
				110				115						120	
Ser	Ser	Thr	Gly	Ala	Ala	Ser	Gly	Ile	Met	Ser	Leu	Ala	Gly	Leu	
				125				130						135	
Val	Leu	Ala	Pro	Phe	Thr	Ala	Gly	Thr	Ser	Leu	Ala	Leu	Thr	Ala	
				140				145						150	
Ala	Gly	Val	Gly	Leu	Gly	Ala	Ala	Ser	Ala	Val	Thr	Gly	Ile	Thr	
				155				160						165	
Thr	Ser	Ile	Val	Glu	His	Ser	Tyr	Thr	Ser	Ser	Ala	Glu	Ala	Glu	
				170				175						180	
Ala	Ser	Arg	Leu	Thr	Ala	Thr	Ser	Ile	Asp	Arg	Leu	Lys	Val	Phe	
				185				190						195	
Lys	Glu	Val	Met	Arg	Asp	Ile	Thr	Pro	Asn	Leu	Leu	Ser	Leu	Leu	
				200				205						210	
Asn	Asn	Tyr	Tyr	Glu	Ala	Thr	Gln	Thr	Ile	Gly	Ser	Glu	Ile	Arg	
				215				220						225	
Ala	Ile	Arg	Gln	Ala	Arg	Ala	Arg	Ala	Arg	Leu	Pro	Val	Thr	Thr	
				230				235						240	
Trp	Arg	Ile	Ser	Ala	Gly	Ser	Gly	Gly	Gln	Ala	Glu	Arg	Thr	Ile	

```
                245                 250                 255
Ala Gly Thr Thr Arg Ala Val Ser Arg Gly Ala Arg Ile Leu Ser
                260                 265                 270
Ala Thr Thr Ser Gly Ile Phe Leu Ala Leu Asp Val Val Asn Leu
                275                 280                 285
Val Tyr Glu Ser Lys His Leu His Glu Gly Ala Lys Ser Ala Ser
                290                 295                 300
Ala Glu Glu Leu Arg Arg Gln Ala Gln Glu Leu Glu Glu Asn Leu
                305                 310                 315
Met Glu Leu Thr Gln Ile Tyr Gln Arg Leu Asn Pro Cys His Thr
                320                 325                 330
His
```

<210> 2
<211> 480
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1292575CD1

<400> 2

```
Met Asn Gly Glu Glu Glu Phe Phe Asp Ala Val Thr Gly Phe Asp
  1               5                  10                  15
Ser Asp Asn Ser Ser Gly Glu Phe Ser Glu Ala Asn Gln Lys Val
                20                  25                  30
Thr Gly Met Ile Asp Leu Asp Thr Ser Lys Asn Asn Arg Ile Gly
                35                  40                  45
Lys Thr Gly Glu Arg Pro Ser Gln Glu Asn Gly Ile Gln Lys His
                50                  55                  60
Arg Thr Ser Leu Pro Ala Pro Met Phe Ser Arg Ser Asp Phe Ser
                65                  70                  75
Val Trp Thr Ile Leu Lys Lys Cys Val Gly Leu Glu Leu Ser Lys
                80                  85                  90
Ile Thr Met Pro Ile Ala Phe Asn Glu Pro Leu Ser Phe Leu Gln
                95                 100                 105
Arg Ile Thr Glu Tyr Met Glu His Val Tyr Leu Ile His Arg Ala
               110                 115                 120
Ser Cys Gln Pro Gln Pro Leu Glu Arg Met Gln Ser Val Ala Ala
               125                 130                 135
Phe Ala Val Ser Ala Val Ala Ser Gln Trp Glu Arg Thr Gly Lys
               140                 145                 150
Pro Phe Asn Pro Leu Leu Gly Glu Thr Tyr Glu Leu Ile Arg Glu
               155                 160                 165
Asp Leu Gly Phe Arg Phe Ile Ser Glu Gln Val Ser His His Pro
               170                 175                 180
Pro Ile Ser Ala Phe His Ser Glu Gly Leu Asn His Asp Phe Leu
               185                 190                 195
Phe His Gly Ser Ile Tyr Pro Lys Leu Lys Phe Trp Gly Lys Ser
               200                 205                 210
Val Glu Ala Glu Pro Arg Gly Thr Ile Thr Leu Glu Leu Leu Lys
               215                 220                 225
His Asn Glu Ala Tyr Thr Trp Thr Asn Pro Thr Cys Cys Val His
               230                 235                 240
Asn Val Ile Ile Gly Lys Leu Trp Ile Glu Gln Tyr Gly Thr Val
               245                 250                 255
Glu Ile Leu Asn His Arg Thr Gly His Lys Cys Val Leu His Phe
               260                 265                 270
Lys Pro Cys Gly Leu Phe Gly Lys Glu Leu His Lys Val Glu Gly
               275                 280                 285
His Ile Gln Asp Lys Asn Lys Lys Lys Leu Phe Met Ile Tyr Gly
               290                 295                 300
Lys Trp Thr Glu Cys Leu Trp Gly Ile Asp Pro Val Ser Tyr Glu
               305                 310                 315
Ser Phe Lys Lys Gln Glu Arg Arg Gly Asp His Leu Arg Lys Ala
               320                 325                 330
```

```
Lys Leu Asp Glu Asp Ser Gly Lys Ala Asp Ser Asp Val Ala Asp
                335                 340                 345
Asp Val Pro Val Ala Gln Glu Thr Val Gln Val Ile Pro Gly Ser
                350                 355                 360
Lys Leu Leu Trp Arg Ile Asn Thr Arg Pro Pro Asn Ser Ala Gln
                365                 370                 375
Met Tyr Asn Phe Thr Ser Phe Thr Val Ser Leu Asn Glu Leu Glu
                380                 385                 390
Thr Gly Met Glu Lys Thr Leu Pro Pro Thr Asp Cys Arg Leu Arg
                395                 400                 405
Pro Asp Ile Arg Gly Met Glu Asn Gly Asn Met Asp Leu Ala Ser
                410                 415                 420
Gln Glu Lys Glu Arg Leu Glu Glu Lys Gln Arg Glu Ala Arg Arg
                425                 430                 435
Glu Arg Ala Lys Glu Glu Ala Glu Trp Gln Thr Arg Trp Phe Tyr
                440                 445                 450
Pro Gly Asn Asn Pro Tyr Thr Gly Thr Pro Asp Trp Leu Tyr Ala
                455                 460                 465
Gly Asp Tyr Phe Glu Arg Asn Phe Ser Asp Cys Pro Asp Ile Tyr
                470                 475                 480


<210> 3
<211> 409
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2454393CD1

<400> 3
Met Ala Thr Ser Val Gly His Arg Cys Leu Gly Leu Leu His Gly
  1               5                  10                  15
Val Ala Pro Trp Arg Ser Ser Leu His Pro Cys Glu Ile Thr Ala
                20                  25                  30
Leu Ser Gln Ser Leu Gln Pro Leu Arg Lys Leu Pro Phe Arg Ala
                35                  40                  45
Phe Arg Thr Asp Ala Arg Lys Ile His Thr Ala Pro Ala Arg Thr
                50                  55                  60
Met Phe Leu Leu Arg Pro Leu Pro Ile Leu Leu Val Thr Gly Gly
                65                  70                  75
Gly Tyr Ala Gly Tyr Arg Gln Tyr Glu Lys Tyr Arg Glu Arg Glu
                80                  85                  90
Leu Glu Lys Leu Gly Leu Glu Ile Pro Pro Lys Leu Ala Gly His
                95                 100                 105
Trp Glu Val Ala Leu Tyr Lys Ser Val Pro Thr Arg Leu Leu Ser
                110                 115                 120
Arg Ala Trp Gly Arg Leu Asn Gln Val Glu Leu Pro His Trp Leu
                125                 130                 135
Arg Arg Pro Val Tyr Ser Leu Tyr Ile Trp Thr Phe Gly Val Asn
                140                 145                 150
Met Lys Glu Ala Ala Val Glu Asp Leu His His Tyr Arg Asn Leu
                155                 160                 165
Ser Glu Phe Phe Arg Arg Lys Leu Lys Pro Gln Ala Arg Pro Val
                170                 175                 180
Cys Gly Leu His Ser Val Ile Ser Pro Ser Asp Gly Arg Ile Leu
                185                 190                 195
Asn Phe Gly Gln Val Lys Asn Cys Glu Val Glu Gln Val Lys Gly
                200                 205                 210
Val Thr Tyr Ser Leu Glu Ser Phe Leu Gly Pro Arg Met Cys Thr
                215                 220                 225
Glu Asp Leu Pro Phe Pro Pro Ala Ala Ser Cys Asp Ser Phe Lys
                230                 235                 240
Asn Gln Leu Val Thr Arg Glu Gly Asn Glu Leu Tyr His Cys Val
                245                 250                 255
Ile Tyr Leu Ala Pro Gly Asp Tyr His Cys Phe His Ser Pro Thr
                260                 265                 270
```

```
Asp Trp Thr Val Ser His Arg Arg His Phe Pro Gly Ser Leu Met
                275                 280                 285
Ser Val Asn Pro Gly Met Ala Arg Trp Ile Lys Glu Leu Phe Cys
                290                 295                 300
His Asn Glu Arg Val Val Leu Thr Gly Asp Trp Lys His Gly Phe
                305                 310                 315
Phe Ser Leu Thr Ala Val Gly Ala Thr Asn Val Gly Ser Ile Arg
                320                 325                 330
Ile Tyr Phe Asp Arg Asp Leu His Thr Asn Ser Pro Arg His Ser
                335                 340                 345
Lys Gly Ser Tyr Asn Asp Phe Ser Phe Val Thr His Thr Asn Arg
                350                 355                 360
Glu Gly Val Pro Met Arg Lys Gly Glu His Leu Gly Glu Phe Asn
                365                 370                 375
Leu Gly Ser Thr Ile Val Leu Ile Phe Glu Ala Pro Lys Asp Phe
                380                 385                 390
Asn Phe Gln Leu Lys Thr Gly Gln Lys Ile Arg Phe Gly Glu Ala
                395                 400                 405
Leu Gly Ser Leu
```

<210> 4
<211> 759
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2766980CD1

<400> 4

```
Met Glu Ser Ser Pro Phe Asn Arg Arg Gln Trp Thr Ser Leu Ser
  1               5                  10                  15
Leu Arg Val Thr Ala Lys Glu Leu Ser Leu Val Asn Lys Asn Lys
                20                  25                  30
Ser Ser Ala Ile Val Glu Ile Phe Ser Lys Tyr Gln Lys Ala Ala
                35                  40                  45
Glu Glu Thr Asn Met Glu Lys Lys Arg Ser Asn Thr Glu Asn Leu
                50                  55                  60
Ser Gln His Phe Arg Lys Gly Thr Leu Thr Val Leu Lys Lys Lys
                65                  70                  75
Trp Glu Asn Pro Gly Leu Gly Ala Glu Ser His Thr Asp Ser Leu
                80                  85                  90
Arg Asn Ser Ser Thr Glu Ile Arg His Arg Ala Asp His Pro Pro
                95                 100                 105
Ala Glu Val Thr Ser His Ala Ala Ser Gly Ala Lys Ala Asp Gln
               110                 115                 120
Glu Glu Gln Ile His Pro Arg Ser Arg Leu Arg Ser Pro Pro Glu
               125                 130                 135
Ala Leu Val Gln Gly Arg Tyr Pro His Ile Lys Asp Gly Glu Asp
               140                 145                 150
Leu Lys Asp His Ser Thr Glu Ser Lys Lys Met Glu Asn Cys Leu
               155                 160                 165
Gly Glu Ser Arg His Glu Val Glu Lys Ser Glu Ile Ser Glu Asn
               170                 175                 180
Thr Asp Ala Ser Gly Lys Ile Glu Lys Tyr Asn Val Pro Leu Asn
               185                 190                 195
Arg Leu Lys Met Met Phe Glu Lys Gly Glu Pro Thr Gln Thr Lys
               200                 205                 210
Ile Leu Arg Ala Gln Ser Arg Ser Ala Ser Gly Arg Lys Ile Ser
               215                 220                 225
Glu Asn Ser Tyr Ser Leu Asp Asp Leu Glu Ile Gly Pro Gly Gln
               230                 235                 240
Leu Ser Ser Ser Thr Phe Asp Ser Glu Lys Asn Glu Ser Arg Arg
               245                 250                 255
Asn Leu Glu Leu Pro Arg Leu Ser Glu Thr Ser Ile Lys Asp Arg
               260                 265                 270
Met Ala Lys Tyr Gln Ala Ala Val Ser Lys Gln Ser Ser Ser Thr
```

```
                275                 280                 285
Asn Tyr Thr Asn Glu Leu Lys Ala Ser Gly Gly Glu Ile Lys Ile
            290                 295                 300
His Lys Met Glu Gln Lys Glu Asn Val Pro Pro Gly Pro Glu Val
        305                 310                 315
Cys Ile Thr His Gln Glu Gly Glu Lys Ile Ser Ala Asn Glu Asn
    320                 325                 330
Ser Leu Ala Val Arg Ser Thr Pro Ala Glu Asp Asp Ser Arg Asp
335                 340                 345
Ser Gln Val Lys Ser Glu Val Gln Gln Pro Val His Pro Lys Pro
            350                 355                 360
Leu Ser Pro Asp Ser Arg Ala Ser Ser Leu Ser Glu Ser Ser Pro
        365                 370                 375
Pro Lys Ala Met Lys Lys Phe Gln Ala Pro Ala Arg Glu Thr Cys
    380                 385                 390
Val Glu Cys Gln Lys Thr Val Tyr Pro Met Glu Arg Leu Leu Ala
395                 400                 405
Asn Gln Gln Val Phe His Ile Ser Cys Phe Arg Cys Ser Tyr Cys
            410                 415                 420
Asn Asn Lys Leu Ser Leu Gly Thr Tyr Ala Ser Leu His Gly Arg
        425                 430                 435
Ile Tyr Cys Lys Pro His Phe Asn Gln Leu Phe Lys Ser Lys Gly
    440                 445                 450
Asn Tyr Asp Glu Gly Phe Gly His Arg Pro His Lys Asp Leu Trp
455                 460                 465
Ala Ser Lys Asn Glu Asn Glu Glu Ile Leu Glu Arg Pro Ala Gln
            470                 475                 480
Leu Ala Asn Ala Arg Glu Thr Pro His Ser Pro Gly Val Glu Asp
        485                 490                 495
Ala Pro Ile Ala Lys Val Gly Val Leu Ala Ala Ser Met Glu Ala
    500                 505                 510
Lys Ala Ser Ser Gln Gln Glu Lys Glu Asp Lys Pro Ala Glu Thr
515                 520                 525
Lys Lys Leu Arg Ile Ala Trp Pro Pro Pro Thr Glu Leu Gly Ser
            530                 535                 540
Ser Gly Ser Ala Leu Glu Glu Gly Ile Lys Met Ser Lys Pro Lys
        545                 550                 555
Trp Pro Pro Glu Asp Glu Ile Ser Lys Pro Glu Val Pro Glu Asp
    560                 565                 570
Val Asp Leu Asp Leu Lys Lys Leu Arg Arg Ser Ser Ser Leu Lys
575                 580                 585
Glu Arg Ser Arg Pro Phe Thr Val Ala Ala Ser Phe Gln Ser Thr
            590                 595                 600
Ser Val Lys Ser Pro Lys Thr Val Ser Pro Pro Ile Arg Lys Gly
        605                 610                 615
Trp Ser Met Ser Glu Gln Ser Glu Glu Ser Val Gly Gly Arg Val
    620                 625                 630
Ala Glu Arg Lys Gln Val Glu Asn Ala Lys Ala Ser Lys Lys Asn
635                 640                 645
Gly Asn Val Gly Lys Thr Thr Trp Gln Asn Lys Glu Ser Lys Gly
            650                 655                 660
Glu Thr Gly Lys Arg Ser Lys Glu Gly His Ser Leu Glu Met Glu
        665                 670                 675
Asn Glu Asn Leu Val Glu Asn Gly Ala Asp Ser Asp Glu Asp Asp
    680                 685                 690
Asn Ser Phe Leu Lys Gln Gln Ser Pro Gln Glu Pro Lys Ser Leu
695                 700                 705
Asn Trp Ser Ser Phe Val Asp Asn Thr Phe Ala Glu Glu Phe Thr
            710                 715                 720
Thr Gln Asn Gln Lys Ser Gln Asp Val Glu Leu Trp Glu Gly Glu
        725                 730                 735
Val Val Lys Glu Leu Ser Val Glu Glu Gln Ile Lys Arg Asn Arg
    740                 745                 750
Tyr Tyr Asp Glu Asp Glu Asp Glu Glu
755
```

<210> 5

<211> 226
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2768356CD1

<400> 5
Met Glu Ile Val Pro Ile Gly Thr Thr His Val Thr Leu Pro Val
1 5 10 15
Phe Gly Asp His Phe Glu Trp Asn Lys Val Thr Ser Cys Ile His
20 25 30
Asn Ile Leu Ser Gly Gln Arg Trp Ile Glu His Tyr Gly Glu Ile
35 40 45
Val Ile Lys Asn Leu His Asp Asp Ser Cys Tyr Cys Lys Val Asn
50 55 60
Phe Ile Lys Ala Lys Tyr Trp Ser Thr Asn Ala His Glu Ile Glu
65 70 75
Gly Thr Val Phe Asp Arg Ser Gly Lys Ala Val His Arg Leu Phe
80 85 90
Gly Lys Trp His Glu Ser Ile Tyr Cys Gly Gly Gly Ser Ser Ser
95 100 105
Ala Cys Val Trp Arg Ala Asn Pro Met Pro Lys Gly Tyr Glu Gln
110 115 120
Tyr Tyr Ser Phe Thr Gln Phe Ala Leu Glu Leu Asn Glu Met Asp
125 130 135
Pro Ser Ser Lys Ser Leu Leu Pro Pro Thr Asp Thr Arg Phe Arg
140 145 150
Pro Asp Gln Arg Phe Leu Glu Glu Gly Asn Leu Glu Glu Ala Glu
155 160 165
Ile Gln Lys Gln Arg Ile Glu Gln Leu Gln Arg Glu Arg Arg Arg
170 175 180
Val Leu Glu Glu Asn His Val Glu His Gln Pro Arg Phe Phe Arg
185 190 195
Lys Ser Asp Asp Asp Ser Trp Val Ser Asn Gly Thr Tyr Leu Glu
200 205 210
Leu Arg Lys Asp Leu Gly Phe Ser Lys Leu Asp His Pro Val Leu
215 220 225
Trp

<210> 6
<211> 500
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5324145CD1

<400> 6
Met Tyr Cys Pro Glu Ser Ala Val Ile Leu Leu Ser Thr Thr Val
1 5 10 15
Leu Glu Asn Val Leu Gln Pro Phe His Phe Arg Ala Gly Thr Met
20 25 30
Ser Lys Leu Pro Lys Phe Glu Ile Glu Leu Pro Ala Ala Pro Lys
35 40 45
Ser Thr Lys Pro Ser Leu Ser Glu Arg Asp Ile Ala Met Ala Thr
50 55 60
Ile Tyr Gly Gln Leu Tyr Val Leu Phe Leu Arg His His Ser Arg
65 70 75
Thr Ser Asn Ser Thr Gly Ala Glu Val Val Leu Tyr His Leu Pro
80 85 90
Arg Glu Gly Ala Cys Lys Lys Met His Ile Leu Lys Leu Asn Arg
95 100 105
Thr Gly Lys Phe Ala Leu Asn Val Val Asp Asn Leu Val Val Val
110 115 120

```
His His Gln Asp Thr Glu Thr Ser Val Ile Phe Asp Ile Lys Leu
                125                 130                 135
Arg Gly Glu Phe Asp Gly Ser Val Thr Phe His His Pro Val Leu
                140                 145                 150
Pro Ala Arg Ser Ile Gln Pro Tyr Gln Ile Pro Ile Thr Gly Pro
                155                 160                 165
Ala Ala Val Thr Ser Gln Ser Pro Val Pro Cys Lys Leu Tyr Ser
                170                 175                 180
Ser Ser Trp Ile Val Phe Gln Pro Asp Ile Ile Ile Ser Ala Ser
                185                 190                 195
Gln Gly Tyr Leu Trp Asn Leu Gln Val Lys Leu Glu Pro Ile Val
                200                 205                 210
Asn Leu Leu Pro Asp Lys Gly Arg Leu Met Asp Phe Leu Leu Gln
                215                 220                 225
Arg Lys Glu Cys Lys Met Val Ile Leu Ser Val Cys Ser Gln Met
                230                 235                 240
Leu Ser Glu Ser Asp Arg Ala Ser Leu Pro Val Ile Ala Thr Val
                245                 250                 255
Phe Asp Lys Leu Asn His Glu Tyr Lys Lys Tyr Leu Asp Ala Glu
                260                 265                 270
Gln Ser Tyr Ala Met Ala Val Glu Ala Gly Gln Ser Arg Ser Ser
                275                 280                 285
Pro Leu Leu Lys Arg Pro Val Arg Thr Gln Ala Val Leu Asp Gln
                290                 295                 300
Ser Asp Val Tyr Thr His Val Leu Ser Ala Phe Val Glu Lys Lys
                305                 310                 315
Glu Met Pro His Lys Phe Val Ile Ala Val Leu Met Glu Tyr Ile
                320                 325                 330
Arg Ser Leu Asn Gln Phe Gln Ile Ala Val Gln His Tyr Leu His
                335                 340                 345
Glu Leu Val Ile Lys Thr Leu Val Gln His Asn Leu Phe Tyr Met
                350                 355                 360
Leu His Gln Phe Leu Gln Tyr His Val Leu Ser Asp Ser Lys Pro
                365                 370                 375
Leu Ala Cys Leu Leu Leu Ser Leu Glu Ser Phe Tyr Pro Pro Ala
                380                 385                 390
His Gln Leu Ser Leu Asp Met Leu Lys Arg Leu Ser Thr Ala Asn
                395                 400                 405
Asp Glu Ile Val Glu Val Leu Leu Ser Lys His Gln Val Leu Ala
                410                 415                 420
Ala Leu Arg Phe Ile Arg Gly Ile Gly Gly His Asp Asn Ile Ser
                425                 430                 435
Ala Arg Lys Phe Leu Asp Ala Ala Lys Gln Thr Glu Asp Asn Met
                440                 445                 450
Leu Phe Tyr Thr Ile Phe Arg Phe Phe Glu Gln Arg Asn Gln Arg
                455                 460                 465
Leu Arg Gly Ser Pro Asn Phe Thr Pro Gly Glu His Cys Glu Glu
                470                 475                 480
His Val Ala Phe Phe Lys Gln Ile Phe Gly Asp Gln Ala Leu Met
                485                 490                 495
Arg Pro Thr Thr Phe
                500
```

<210> 7
<211> 272
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1004646CD1

<400> 7

```
Met Ser Cys His Asn Cys Ser Asp Pro Gln Val Leu Cys Ser Ser
  1               5                  10                  15
Gly Gln Leu Phe Leu Gln Pro Leu Trp Asp His Leu Arg Ser Trp
                20                  25                  30
```

--

```
Glu Ala Leu Leu Gln Ser Pro Phe Phe Pro Val Ile Phe Ser Ile
                35                  40                  45
Thr Thr Tyr Val Gly Phe Cys Leu Pro Phe Val Val Leu Asp Ile
                50                  55                  60
Leu Cys Ser Trp Val Pro Ala Leu Arg Arg Tyr Lys Ile His Pro
                65                  70                  75
Asp Phe Ser Pro Ser Ala Gln Gln Leu Leu Pro Cys Leu Gly Gln
                80                  85                  90
Thr Leu Tyr Gln His Val Met Phe Val Phe Pro Val Thr Leu Leu
                95                 100                 105
His Trp Ala Arg Ser Pro Ala Leu Leu Pro His Glu Ala Pro Glu
               110                 115                 120
Leu Leu Leu Leu Leu His His Ile Leu Phe Cys Leu Leu Leu Phe
               125                 130                 135
Asp Met Glu Phe Phe Val Trp His Leu Leu His His Lys Val Pro
               140                 145                 150
Trp Leu Tyr Arg Thr Phe His Lys Val His His Gln Asn Ser Ser
               155                 160                 165
Ser Phe Ala Leu Ala Thr Gln Tyr Met Ser Val Trp Glu Leu Phe
               170                 175                 180
Ser Leu Gly Phe Phe Asp Met Met Asn Val Thr Leu Leu Gly Cys
               185                 190                 195
His Pro Leu Thr Thr Leu Thr Phe His Val Val Asn Ile Trp Leu
               200                 205                 210
Ser Val Glu Asp His Ser Gly Tyr Asn Phe Pro Trp Ser Thr His
               215                 220                 225
Arg Leu Val Pro Phe Gly Trp Tyr Gly Gly Val Val His His Asp
               230                 235                 240
Leu His His Ser His Phe Asn Cys Asn Phe Ala Pro Tyr Phe Thr
               245                 250                 255
His Trp Asp Lys Ile Leu Gly Thr Leu Arg Thr Ala Ser Val Pro
               260                 265                 270
Ala Arg
```

<210> 8
<211> 282
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1802851CD1

<400> 8

```
Met Ser Gly Gly Trp Met Ala Gln Val Gly Ala Trp Arg Thr Gly
  1               5                  10                  15
Ala Leu Gly Leu Ala Leu Leu Leu Leu Leu Gly Leu Gly Leu Gly
                20                  25                  30
Leu Glu Ala Ala Ala Ser Pro Leu Ser Thr Pro Thr Ser Ala Gln
                35                  40                  45
Ala Ala Gly Pro Ser Ser Gly Ser Cys Pro Pro Thr Lys Phe Gln
                50                  55                  60
Cys Arg Thr Ser Gly Leu Cys Val Pro Leu Thr Trp Arg Cys Asp
                65                  70                  75
Arg Asp Leu Asp Cys Ser Asp Gly Ser Asp Glu Glu Glu Cys Arg
                80                  85                  90
Ile Glu Pro Cys Thr Gln Lys Gly Gln Cys Pro Pro Pro Pro Gly
                95                 100                 105
Leu Pro Cys Pro Cys Thr Gly Val Ser Asp Cys Ser Gly Gly Thr
               110                 115                 120
Asp Lys Lys Leu Arg Asn Cys Ser Arg Leu Ala Cys Leu Ala Gly
               125                 130                 135
Glu Leu Arg Cys Thr Leu Ser Asp Asp Cys Ile Pro Leu Thr Trp
               140                 145                 150
Arg Cys Asp Gly His Pro Asp Cys Pro Asp Ser Ser Asp Glu Leu
               155                 160                 165
Gly Cys Gly Thr Asn Glu Ile Leu Pro Glu Gly Asp Ala Thr Thr
```

```
                  170                 175                 180
Met Gly Pro Pro Val Thr Leu Glu Ser Val Thr Ser Leu Arg Asn
                  185                 190                 195
Ala Thr Thr Met Gly Pro Pro Val Thr Leu Glu Ser Val Pro Ser
                  200                 205                 210
Val Gly Asn Ala Thr Ser Ser Ser Ala Gly Asp Gln Ser Gly Ser
                  215                 220                 225
Pro Thr Ala Tyr Gly Val Ile Ala Ala Ala Ala Val Leu Ser Ala
                  230                 235                 240
Ser Leu Val Thr Ala Thr Leu Leu Leu Leu Ser Trp Leu Arg Ala
                  245                 250                 255
Gln Glu Arg Leu Arg Pro Leu Gly Leu Leu Val Ala Met Lys Glu
                  260                 265                 270
Ser Leu Leu Leu Ser Glu Gln Lys Thr Ser Leu Pro
                  275                 280
```

<210> 9
<211> 437
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2764333CD1

<400> 9

```
Met Ser Glu Glu Lys Asp Cys Gly Gly Gly Asp Ala Leu Ser Asn
  1                 5                  10                  15
Gly Ile Lys Lys His Arg Thr Ser Leu Pro Ser Pro Met Phe Ser
                 20                  25                  30
Arg Asn Asp Phe Ser Ile Trp Ser Ile Leu Arg Lys Cys Ile Gly
                 35                  40                  45
Met Glu Leu Ser Lys Ile Thr Met Pro Val Ile Phe Asn Glu Pro
                 50                  55                  60
Leu Ser Phe Leu Gln Arg Leu Thr Glu Tyr Met Glu His Thr Tyr
                 65                  70                  75
Leu Ile His Lys Ala Ser Ser Leu Ser Asp Pro Val Glu Arg Met
                 80                  85                  90
Gln Cys Val Ala Ala Phe Ala Val Ser Ala Val Ala Ser Gln Trp
                 95                 100                 105
Glu Arg Thr Gly Lys Pro Phe Asn Pro Leu Leu Gly Glu Thr Tyr
                110                 115                 120
Glu Leu Val Arg Asp Asp Leu Gly Phe Arg Leu Ile Ser Glu Gln
                125                 130                 135
Val Ser His His Pro Pro Ile Ser Ala Phe His Ala Glu Gly Leu
                140                 145                 150
Asn Asn Asp Phe Ile Phe His Gly Ser Ile Tyr Pro Lys Leu Lys
                155                 160                 165
Phe Trp Gly Lys Ser Val Glu Ala Glu Pro Lys Gly Thr Ile Thr
                170                 175                 180
Leu Glu Leu Leu Glu His Asn Glu Ala Tyr Thr Trp Thr Asn Pro
                185                 190                 195
Thr Cys Cys Val His Asn Ile Ile Val Gly Lys Leu Trp Ile Glu
                200                 205                 210
Gln Tyr Gly Asn Val Glu Ile Ile Asn His Lys Thr Gly Asp Lys
                215                 220                 225
Cys Val Leu Asn Phe Lys Pro Cys Gly Leu Phe Gly Lys Glu Leu
                230                 235                 240
His Lys Val Glu Gly Tyr Ile Gln Asp Lys Ser Lys Lys Lys Leu
                245                 250                 255
Cys Ala Leu Tyr Gly Lys Trp Thr Glu Cys Leu Tyr Ser Val Asp
                260                 265                 270
Pro Ala Thr Phe Asp Ala Tyr Lys Lys Asn Asp Lys Lys Asn Thr
                275                 280                 285
Glu Glu Lys Lys Asn Ser Lys Gln Met Ser Thr Ser Glu Glu Leu
                290                 295                 300
Asp Glu Met Pro Val Pro Asp Ser Glu Ser Val Phe Ile Ile Pro
```

```
                305                 310                 315
Gly Ser Val Leu Leu Trp Arg Ile Ala Pro Arg Pro Pro Asn Ser
                320                 325                 330
Ala Gln Met Tyr Asn Phe Thr Ser Phe Ala Met Val Leu Asn Glu
                335                 340                 345
Val Asp Lys Asp Met Glu Ser Val Ile Pro Lys Thr Asp Cys Arg
                350                 355                 360
Leu Arg Pro Asp Ile Arg Ala Met Glu Asn Gly Glu Ile Asp Gln
                365                 370                 375
Ala Ser Glu Glu Lys Lys Arg Leu Glu Glu Lys Gln Arg Ala Ala
                380                 385                 390
Arg Lys Asn Arg Ser Lys Ser Glu Glu Asp Trp Lys Thr Arg Trp
                395                 400                 405
Phe His Gln Gly Pro Asn Pro Tyr Asn Gly Ala Gln Asp Trp Ile
                410                 415                 420
Tyr Ser Gly Ser Tyr Trp Asp Arg Asn Tyr Phe Asn Leu Pro Asp
                425                 430                 435
Ile Tyr
```

<210> 10
<211> 427
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2798021CD1

<400> 10

```
Met Arg Gln Ala Ala Ala Asp Ala Lys Pro Glu Ser Leu Met Lys
  1               5                  10                  15
Arg Leu Glu Glu Glu Ile Lys Phe Asn Leu Tyr Met Val Thr Glu
                20                  25                  30
Lys Phe Pro Lys Glu Leu Glu Asn Lys Lys Lys Glu Leu His Phe
                35                  40                  45
Leu Gln Lys Val Val Ser Glu Pro Ala Met Gly His Ser Asp Leu
                50                  55                  60
Leu Glu Leu Glu Ser Lys Ile Asn Glu Ile Asn Thr Glu Ile Asn
                65                  70                  75
Gln Leu Ile Glu Lys Lys Met Met Arg Asn Glu Pro Ile Glu Gly
                80                  85                  90
Lys Leu Ser Leu Tyr Arg Gln Gln Ala Ser Ile Ile Ser Arg Lys
                95                 100                 105
Lys Glu Ala Lys Ala Glu Glu Leu Gln Glu Ala Lys Glu Lys Leu
               110                 115                 120
Ala Ser Leu Glu Arg Glu Ala Ser Val Lys Arg Asn Gln Thr Arg
               125                 130                 135
Glu Phe Asp Gly Thr Glu Val Leu Lys Gly Asp Glu Phe Lys Arg
               140                 145                 150
Tyr Val Asn Lys Leu Arg Ser Lys Ser Thr Val Phe Lys Lys Lys
               155                 160                 165
His Gln Ile Ile Ala Glu Leu Lys Ala Glu Phe Gly Leu Leu Gln
               170                 175                 180
Arg Thr Glu Glu Leu Leu Lys Gln Arg His Glu Asn Ile Gln Gln
               185                 190                 195
Gln Leu Gln Thr Met Glu Glu Lys Lys Gly Ile Ser Gly Tyr Ser
               200                 205                 210
Tyr Thr Gln Glu Glu Leu Glu Arg Val Ser Ala Leu Lys Ser Glu
               215                 220                 225
Val Asp Glu Met Lys Gly Arg Thr Leu Asp Asp Met Ser Glu Met
               230                 235                 240
Val Lys Lys Leu Tyr Ser Leu Val Ser Glu Lys Lys Ser Ala Leu
               245                 250                 255
Ala Ser Val Ile Lys Glu Leu Arg Gln Leu Arg Gln Lys Tyr Gln
               260                 265                 270
Glu Leu Thr Gln Glu Cys Asp Glu Lys Lys Ser Gln Tyr Asp Ser
               275                 280                 285
```

```
Cys Ala Ala Gly Leu Glu Ser Asn Arg Ser Lys Leu Glu Gln Glu
                290                 295                 300
Val Arg Arg Leu Arg Glu Glu Cys Leu Gln Glu Glu Ser Arg Tyr
                305                 310                 315
His Tyr Thr Asn Cys Met Ile Lys Asn Leu Glu Val Gln Leu Arg
                320                 325                 330
Arg Ala Thr Asp Glu Met Lys Ala Tyr Ile Ser Ser Asp Gln Gln
                335                 340                 345
Glu Lys Arg Lys Ala Ile Arg Glu Gln Tyr Thr Lys Asn Thr Ala
                350                 355                 360
Glu Gln Glu Asn Leu Gly Lys Lys Leu Arg Glu Lys Gln Lys Val
                365                 370                 375
Ile Arg Glu Ser His Gly Pro Asn Met Lys Gln Ala Lys Met Trp
                380                 385                 390
Arg Asp Leu Glu Gln Leu Met Glu Cys Lys Lys Gln Cys Phe Leu
                395                 400                 405
Lys Gln Gln Ser Gln Thr Ser Ile Gly Gln Val Ile Gln Glu Gly
                410                 415                 420
Gly Glu Asp Arg Leu Ile Leu
                425
```

<210> 11
<211> 564
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3335404CD1

<400> 11

```
Met Asp Ser Arg Tyr Asn Ser Thr Ala Gly Ile Gly Asp Leu Asn
  1                 5                  10                  15
Gln Leu Ser Ala Ala Ile Pro Ala Thr Arg Val Glu Val Ser Val
                 20                  25                  30
Ser Cys Arg Asn Leu Leu Asp Arg Asp Thr Phe Ser Lys Ser Asp
                 35                  40                  45
Pro Ile Cys Val Leu Tyr Val Gln Gly Val Gly Asn Lys Glu Trp
                 50                  55                  60
Arg Glu Phe Gly Arg Thr Glu Val Ile Asp Asn Thr Leu Asn Pro
                 65                  70                  75
Asp Phe Val Arg Lys Phe Ile Leu Asp Tyr Phe Phe Glu Glu Arg
                 80                  85                  90
Glu Asn Leu Arg Phe Asp Leu Tyr Asp Val Asp Ser Lys Ser Pro
                 95                 100                 105
Asn Leu Ser Lys His Asp Phe Leu Gly Gln Val Phe Cys Thr Leu
                110                 115                 120
Gly Glu Ile Val Gly Ser Gln Gly Ser Arg Leu Glu Lys Pro Ile
                125                 130                 135
Val Gly Ile Pro Gly Lys Lys Cys Gly Thr Ile Ile Leu Thr Ala
                140                 145                 150
Glu Glu Leu Asn Cys Cys Arg Asp Ala Val Leu Met Gln Phe Cys
                155                 160                 165
Ala Asn Lys Leu Asp Lys Lys Asp Phe Phe Gly Lys Ser Asp Pro
                170                 175                 180
Phe Leu Val Phe Tyr Arg Ser Asn Glu Asp Gly Ser Phe Thr Ile
                185                 190                 195
Cys His Lys Thr Glu Val Val Lys Asn Thr Leu Asn Pro Val Trp
                200                 205                 210
Gln Ala Phe Lys Ile Ser Val Arg Ala Leu Cys Asn Gly Asp Tyr
                215                 220                 225
Asp Arg Thr Ile Lys Val Glu Val Tyr Asp Trp Asp Arg Asp Gly
                230                 235                 240
Ser His Asp Phe Ile Gly Glu Phe Thr Thr Ser Tyr Arg Glu Leu
                245                 250                 255
Ser Arg Gly Gln Ser Gln Phe Asn Val Tyr Glu Val Val Asn Pro
                260                 265                 270
```

```
Lys Lys Lys Gly Lys Lys Lys Lys Tyr Thr Asn Ser Gly Thr Val
                275                 280                 285
Thr Leu Leu Ser Phe Leu Val Glu Thr Glu Val Ser Phe Leu Asp
                290                 295                 300
Tyr Ile Lys Gly Gly Thr Gln Ile Asn Phe Thr Val Ala Ile Asp
                305                 310                 315
Phe Thr Ala Ser Asn Gly Asn Pro Ala Gln Pro Thr Ser Leu His
                320                 325                 330
Tyr Met Asn Pro Tyr Gln Leu Asn Ala Tyr Gly Met Ala Leu Lys
                335                 340                 345
Ala Val Gly Glu Ile Val Gln Asp Tyr Asp Ser Asp Lys Met Phe
                350                 355                 360
Pro Ala Leu Gly Phe Gly Ala Lys Leu Pro Pro Asp Gly Arg Ile
                365                 370                 375
Ser His Glu Phe Ala Leu Asn Gly Asn Pro Gln Asn Pro Tyr Cys
                380                 385                 390
Asp Gly Ile Glu Gly Val Met Glu Ala Tyr Tyr Arg Ser Leu Lys
                395                 400                 405
Ser Val Gln Leu Tyr Gly Pro Thr Asn Phe Ala Pro Val Ile Asn
                410                 415                 420
His Val Ala Arg Tyr Ala Ser Ser Val Lys Asp Gly Ser Gln Tyr
                425                 430                 435
Phe Val Leu Leu Ile Val Thr Asp Gly Val Ile Ser Asp Met Ala
                440                 445                 450
Gln Thr Lys Glu Ser Ile Val Asn Ala Ser Lys Leu Pro Met Ser
                455                 460                 465
Ile Ile Ile Val Gly Val Gly Pro Ala Glu Phe Asp Ala Met Val
                470                 475                 480
Glu Leu Asp Gly Asp Asp Val Arg Val Ser Ser Arg Gly Lys Tyr
                485                 490                 495
Ala Glu Arg Asp Ile Val Gln Phe Val Pro Phe Arg Asp Tyr Ile
                500                 505                 510
Asp Arg Ser Gly Asn His Ile Leu Ser Met Ala Arg Leu Ala Lys
                515                 520                 525
Asp Val Leu Ala Glu Ile Pro Glu Gln Phe Leu Ser Tyr Met Arg
                530                 535                 540
Ala Arg Gly Ile Lys Pro Ser Pro Ala Pro Pro Pro Tyr Thr Pro
                545                 550                 555
Pro Thr His Val Leu Gln Thr Gln Ile
                560
```

<210> 12
<211> 297
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3735780CD1

<400> 12

```
Met Met Asp Ser Glu Ala His Glu Lys Arg Pro Pro Ile Leu Thr
  1                 5                  10                  15
Ser Ser Lys Gln Asp Ile Ser Pro His Ile Thr Asn Val Gly Glu
                 20                  25                  30
Met Lys His Tyr Leu Cys Gly Cys Cys Ala Ala Phe Asn Asn Val
                 35                  40                  45
Ala Ile Thr Phe Pro Ile Gln Lys Val Leu Phe Arg Gln Gln Leu
                 50                  55                  60
Tyr Gly Ile Lys Thr Arg Asp Ala Ile Leu Gln Leu Arg Arg Asp
                 65                  70                  75
Gly Phe Arg Asn Leu Tyr Arg Gly Ile Leu Pro Pro Leu Met Gln
                 80                  85                  90
Lys Thr Thr Thr Leu Ala Leu Met Phe Gly Leu Tyr Glu Asp Leu
                 95                 100                 105
Ser Cys Leu Leu His Lys His Val Ser Ala Pro Glu Phe Ala Thr
                110                 115                 120
```

```
Ser Gly Val Ala Ala Val Leu Ala Gly Thr Thr Glu Ala Ile Phe
                125                 130                 135
Thr Pro Leu Glu Arg Val Gln Thr Leu Leu Gln Asp His Lys His
                140                 145                 150
His Asp Lys Phe Thr Asn Thr Tyr Gln Ala Phe Lys Ala Leu Lys
                155                 160                 165
Cys His Gly Ile Gly Glu Tyr Tyr Arg Gly Leu Val Pro Ile Leu
                170                 175                 180
Phe Arg Asn Gly Leu Ser Asn Val Leu Phe Phe Gly Leu Arg Gly
                185                 190                 195
Pro Ile Lys Glu His Leu Pro Thr Ala Thr Thr His Ser Ala His
                200                 205                 210
Leu Val Asn Asp Phe Ile Cys Gly Gly Leu Leu Gly Ala Met Leu
                215                 220                 225
Gly Phe Leu Phe Phe Pro Ile Asn Val Val Lys Thr Arg Ile Gln
                230                 235                 240
Ser Gln Ile Gly Gly Glu Phe Gln Ser Phe Pro Lys Val Phe Gln
                245                 250                 255
Lys Ile Trp Leu Glu Arg Asp Arg Lys Leu Ile Asn Leu Phe Arg
                260                 265                 270
Gly Ala His Leu Asn Tyr His Arg Ser Leu Ile Ser Trp Gly Ile
                275                 280                 285
Ile Asn Ala Thr Tyr Glu Phe Leu Leu Lys Val Ile
                290                 295
```

<210> 13
<211> 2174
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 161190CB1

<400> 13

```
ggatccacac agctcagaac agctggatct tgctcacact ctttcaagag aagcttcctt  60
ggacaaaagg accctgcctt ggtgtgagag tgagggcaga gggagctgga gcaagtagaa 120
tttctctaaa taccagctgg ctggggccca ggagattaaa aaacaccggg ctagggttaa 180
gaaaaaaaac gaacccttcc agtcaggtca gtgactggag agctccaagg aaagtctctc 240
agtgacctgg ctgctggcac catggactca gaaaagaaac gctttactga agaggccacc 300
aaatacttcc gggagagagt cagcccagtg catctgcaaa tcctgctgac taacaatgaa 360
gcctggaaga gattcgtgac tgcggctgaa ttgcccaggg atgaggcaga tgctctctac 420
gaagctctga agaagcttag aacatatgca gctattgagg acgaatatgt gcagcagaaa 480
gatgagcagt ttagggaatg gtttttgaaa gagtttcccc aagtcaagag gaagatccag 540
gagtccatag aaaagcttcg tgcccttgca aatggtattg aagaggtcca cagaggctgc 600
accatctcca atgtggtgtc cagctccact ggcgctgcct ctggcatcat gtcccttgct 660
ggtcttgttt tggcaccatt tacagcaggg acgagtctgg cccttactgc agctggggta 720
gggctgggag cagcgtctgc tgtgactggg atcaccacca gcatcgtgga gcactcatac 780
acatcatcag cagaagctga agccagcagg ctgactgcaa ccagcattga ccgattgaag 840
gtatttaagg aagttatgcg tgacatcaca cccaacttac tttcccttct taataattat 900
tacgaagcca cacaaaccat tgggagtgaa atccgtgcca tcaggcaagc cagagccagg 960
gcccgactcc ctgtgaccac ctggcgaatc tcagctggaa gtggaggtca agcagagaga 1020
acgattgcag gcaccacccg ggcagtgagc agaggagccc ggatcctgag tgcgaccact 1080
tcaggcatct tccttgcact ggatgtggtc aaccttgtat acgagtcaaa gcacttgcat 1140
gaggggcaa agtctgcatc tgctgaggag ctgaggcggc aggctcagga gctggaggag 1200
aatctaatgg agctcactca gatctatcag cgtctgaatc catgccatac ccactgaccc 1260
cagaccagtg cagccagcag gggaggtgag ccatacacag gccacgacaa aatgcaggca 1320
ttttattagg gggataaaga gggcaaggta aagtttatgg agctgagtgt tagtgacttt 1380
ggcatttctg tagctgagca cagcagggga ggggttaatg cagatggcaa gtgcaccaag 1440
gagaaggcag gaatgctgga gcctggaata agggaagaga ggggactgga gagtgtgggg 1500
aataggaaga agaaatttcc tttagactaa cgaatatatt ggggggagga atagagggga 1560
ggtgtgcagg aaccagcaat gagaaggcca ggaaaagaaa gagctgaaaa tgcagaaagc 1620
cgaagagtta gaacttttgg atacagcaga agaaacagcg gctccactac cgacctgccc 1680
ccggttcgat gtccttccaa gaatgaagtc tttccctggt gatggtcccc tgccctgtct 1740
ttccagcatc cactctgtct tgtcctcctg gaagtgtatc tcagtcagcc agtggcttct 1800
tgatgatggc ggtggaggtg gtggttgtag tgtgatggat ccccttttagg ttatttaggg 1860
gtatatgtcc cctgcttgaa ccctgaaggc caggtaatga gccatggcca ttgtccccag 1920
```

ctgaggacca ggtgtctcta aaaacccaaa catcctggag agtatgcgag aacctaccaa 1980
gaaaaacagt ctcattactc atatacagca ggcaaagaga cagaaaatta actgaaaagc 2040
agtttagaga ctgggggagg ccggatctct agagccatcc tgctgagtgc cctgtgtgta 2100
agtcctaata aactcaccta ctcaccaaaa aaaaaaacga aaaactaaag aacaggagaa 2160
aaaaagggga gggc 2174

<210> 14
<211> 2620
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1292575CB1

<400> 14
agtgagctcg gccggcaacc gagggacccg cgtccagatc ttcagtgtct attggatttt 60
tccaagagaa agtttgtaaa attccttaca ctgtagatgt ggatcagata cgatgattca 120
gtagaagagc acatgtcagg ggcagtggag gctggctgct gaaggatgaa cggagaggaa 180
gaattctttg atgccgtcac aggctttgat tctgataact cttctgggga attttcagag 240
gcaaatcaga aagtcacggg aatgattgac ttagacacca gcaaaaataa taggattggg 300
aaaactgggg agaggccctc tcaagagaac ggaattcaga aacacaggac atcgctgccg 360
gctcccatgt tcagcagaag cgacttcagc gtgtggacca tcctgaagaa gtgtgttggc 420
ctggagctgt ccaagatcac gatgccaatc gccttcaacg agcctctgag cttcttgcag 480
cggatcacgg agtacatgga gcacgtgtac ctcatccaca gggcctcctg ccagccccag 540
cccctggaga ggatgcagtc tgtggctgct tttgctgttt cggctgtggc ttcccagtgg 600
gagaggaccg gcaaaccatt taatccactc ttgggagaaa cgtatgaatt aatcagggaa 660
gatttaggat tcagatttat atcggaacag gtcagtcacc acccccccat cagtgcgttc 720
cactcggaag gtctcaacca tgacttcctg ttccatggct ccatctaccc caagctcaag 780
ttctggggca aaagcgtgga ggcggagccc cgaggcacca tcaccctgga gctgctcaaa 840
cataatgaag cctacacctg gaccaacccc acctgctgcg tccacaacgt catcatcggg 900
aagctgtgga tagagcagta tgggacagtg gagattttaa accacagaac tggacataag 960
tgtgtgcttc actttaaacc gtgtggatta tttggaaaag aacttcacaa ggtggaagga 1020
cacattcaag acaaaaacaa aaagaagctc tttatgatct atggcaaatg gacggaatgt 1080
ttgtggggca tagatcctgt ttcgtatgaa tccttcaaga agcaggagag gagaggtgac 1140
cacctgagaa aggccaagct ggatgaagac tccgggaagg ctgacagcga cgtggctgac 1200
gacgtgcctg tggcccagga gaccgtgcag gtcattcctg gcagcaagct gctctggagg 1260
atcaacaccc ggccccccaa ctctgcccag atgtataatt tcaccagttt cactgtgagc 1320
ctcaacgagc tggagacagg catggagaag accctgccac ccacggactg ccgcctgcgc 1380
cctgacatcc gcggcatgga gaatggcaac atggatctgg ccagccagga gaaggagcgg 1440
ctggaggaga agcagagaga agcacggagg gagcgggcca aggaggaggc agagtggcag 1500
acgaggtggt tctacccagg caataacccc tacactggga cccccgactg gttgtatgca 1560
ggggattact ttgagcggaa tttctccgac tgcccagata tctactgagg gcctggaggg 1620
gcctggggcc cgggaccgga ggctgacgag gctggacttc ctcgagtggc cactgtgagc 1680
ctcgtcacag cagaaaccaa cttttctaac gactgagttc gcggagatag catcatccct 1740
gatcaaggat gtaattctaa ttaactgttg attgccaaac atttcactct gctgtgccgt 1800
ctcttcataa agcttcactt gggatcatcg tcttcattaa ggtttcaaca gggaaattct 1860
tcacggcgcc cttttatgtg gcagaaatca gctggggctt gtttagcttc cagcacactc 1920
tcagtcatag catgtgtagc taaaggaagt aatgggaagg ggttcatgtt ctctttataa 1980
tgcagtggca aaaggttctg aaagcctttt aaactcgaac cagtgggggа aagatggatc 2040
ttgaagctaa tcctgcagag agttttatag aggccaggga ttgccttcta aattatgata 2100
aaacagaagt gaagagtttc agagcatcag attgagtgaa aagttgtcag attctgtatt 2160
ttttaacaat cttcaataat gtaaagatta cttttaaaat atttaagtta aaactacttg 2220
aatagtattt tgctgaagag caagatatgc attaatcacc ggttttatac tgtccaaaat 2280
gaagcatccc cgtgacaaac cagagtgggc agaagcatcg agagcgtgac aggaaatccc 2340
aagactgctt ccgcctcaga ggcgtcccgg ctgcgattcg ctgccctgtt gtcagtgagg 2400
cctggctgtc accgcacacc gcgtccgtgt ctccaggggg ttcctttctt ctcacacgtc 2460
gcgtgtaccc atagcactct tgtgtttctg tttttcccag tatgcatgtt taaaatagaa 2520
gtgacaagaa tcacatccgg ttgtgtcctg tgggagggtc agaggcagaa tctacttaca 2580
gtggtgtaat taaagttatt taaccaaaaa aaaaaaaaaa 2620

<210> 15
<211> 2426
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2454393CB1

<400> 15
gaaacactag acatgatacg aattcgagct cgtacccttc gaggcacgct gagaaggagc 60
agacaagatg gcgacgtccg tggggcaccg atgtctggga ttactgcacg gggtcgcgcc 120
gtggcggagc agcctccatc cctgtgagat cactgccctg agccaatccc tacagccctt 180
acggaagctg ccttttagag cctttcgcac agatgccaga aaaatccaca ctgcccctgc 240
ccgaaccatg ttcctgctgc gtcccctgcc cattctgttg gtgacaggcg gcgggtatgc 300
agggtaccgg cagtatgaga agtacaggga gcgagagctg gagaagctgg gattggagat 360
tccacccaaa cttgctggtc actgggaggt ggctttgtac aagtcagtgc caacgcgctt 420
gctgtcacgg gcctggggtc gcctcaatca ggtggagctg ccacactggc tgcgcaggcc 480
cgtctacagc ctgtacatct ggacgtttgg ggtgaacatg aaagaggccg ctgtggagga 540
cctgcatcac taccgcaacc tcagcgagtt cttccggcgc aagctgaagc cgcaggcccg 600
gcctgtctgt ggcctgcaca gcgtgattag cccatcggat ggaaggatcc tcaactttgg 660
gcaggtgaag aactgtgagg tggagcaggt aaaggggtc acctactccc tggagtcgtt 720
cctgggcccg cgtatgtgca cagaggacct gcccttccca ccagccgcgt cgtgtgactc 780
cttcaagaac cagctggtca cccgggaagg gaatgagctc tatcactgtg tcatctacct 840
ggcccctggg gactaccact gcttccactc ccccaccgac tggactgtgt cccaccggcg 900
ccacttccca ggctccctga tgtcagtgaa ccctggcatg gctcgctgga tcaaagagct 960
cttctgccat aacgagcggg tggtcctgac gggggactgg aaacatggct tcttctcact 1020
gacagctgtg ggggccacca acgtgggctc cattcgcatc tactttgacc gggacctgca 1080
cacaaacagc ccaaggcaca gcaagggctc ctacaatgac ttcagcttcg tgacgcacac 1140
caatagagag ggcgtcccca tgcgtaaggg cgagcacctg ggcgagttca acctgggctc 1200
caccatcgtg ctcatcttcg aggcccccaa ggacttcaat ttccagctga aaacaggaca 1260
gaaaatccgc tttggggaag ccctgggctc gctctagagt ctctttcctg attatggctg 1320
ctaagggatc ttttccaaac agagtgaggg tcttttcaag agggaggccc atgaggccat 1380
ccaggtaagg gcctgcctca gcgtggttgg gagtctgacc aggtaggact tgaatgattc 1440
ggctaccacc tgttccagag gtgcagacaa gaggtggcga gagcccccat catgcccctc 1500
aacctatccc gttccttctg cctacaaata aaaagtgcag gctggaatga tctcagtcac 1560
atttggatct ttttaaacac tgtatagacg gaagagcctg cattcctgac cgaaccttca 1620
gttggtctcg gttgtcgttt tttcttgctg ctcctccccc catcacctga gctgtttttct 1680
gttggccccc tttgtttttt ggccttaacg ctcctgctgc acagggtgag gtacctcctt 1740
ggcacagact gtggatgcct ctcccccagc agagccacac agccttcgtg acaactgctt 1800
tccgttccca cattcacctc atcctgctct ttagaaaaag cagtctttgt gcttgtggct 1860
gaacgcatca ccctggactc tgctagtgtc ttctgaggac actgatgaca ctgattaatg 1920
atacagacct ttgcaggacc tgatgagtga cccttctgga gctggccagg tcctctgcag 1980
caggcaagac caatcaatca ctgaacctgc ctcatggcac cagagtgaac agggcaggca 2040
ggtagtaggc ccagctgggg aaatgggaga gttcctgtcc ccctccacat atccctacat 2100
gaaatatggg aaagttgctg ctattgattc agggtctgtc ttggaggcag aggacccttg 2160
gtggatagtt ggtcagtgcc tggaaaacct gtcccagttt atcaggaacg caggcctggg 2220
gagcccccag tggcggggac agggccagat ttcatgttga ccctggggat gctgtgaatt 2280
tctcctgcag gagagacatc attgaatttt ttcaactgta tcagtagcac agtatttttg 2340
tatgaaaagt gggagacttc tgaacagtaa ttcatttaat tgcaaagcat tttgaaataa 2400
aaaaaatcaa acttaaaaaa aaaaaa 2426

<210> 16
<211> 3705
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2766980CB1

<400> 16
ggccgcagga gcagtaggtg ttagcagctt ggtcgcgaca ggtgcgctag gtagagcgcc 60
gggacctgtg acagggctgg tagcagcgca gaggaaaggc ggcttttagc caggtatttc 120
agtgtctgta gacaagatgg aatcatctcc atttaataga cggcaatgga cctcactatc 180
attgagggta acagccaaag aactttctct tgtcaacaag aacaagtcat cggctattgt 240
ggaaatattc tccaagtacc agaaagcagc tgaagaaaca aacatggaga agaagagaag 300
taacaccgaa aatctctccc agcactttag aaaggggacc ctgactgtgt taaagaagaa 360
gtgggagaac ccagggctgg gagcagagtc tcacacagac tctctacgga acagcagcac 420
tgagattagg cacagagcag accatcctcc tgctgaagtg acaagccacg ctgcttctgg 480
agccaaagct gaccaagaag aacaaatcca ccccagatct agactcaggt cacctcctga 540
agccctcgtt cagggtcgat atccccacat caaggacggt gaggatctta aagaccactc 600

aacagaaagt aaaaaaatgg aaaattgtct aggagaatcc aggcatgaag tagaaaaatc 660
agaaatcagt gaaaacacag atgcttcggg caaaatagag aaatataatg ttccgctgaa 720
caggcttaag atgatgtttg agaaaggtga accaactcaa actaagattc tccgggccca 780
aagccgaagt gcaagtggaa ggaagatctc tgaaaacagc tattctctag atgacctgga 840
aataggccca ggtcagttgt catcttctac atttgactcg gagaaaaatg agagtagacg 900
aaatctggaa cttccacgcc tctcagaaac ctctataaag gatcgaatgg ccaagtacca 960
ggcagctgtg tccaaacaaa gcagctcaac caactataca aatgagctga aagccagtgg 1020
tggcgaaatc aaaattcata aaatggagca aaaggagaat gtgcccccag gtcctgaggt 1080
ctgcatcacc catcaggaag gggaaaagat ttctgcaaat gagaatagcc tggcagtccg 1140
ttccaccccc gccgaagatg actcccgtga ctcccaggtt aagagtgagg ttcaacagcc 1200
tgtccatccc aagccactaa gtccagattc cagagcctcc agtctttctg aaagttctcc 1260
tcccaaagca atgaagaagt ttcaggcacc tgcaagagag acctgcgtgg aatgtcagaa 1320
gacagtctat ccaatggagc gtctcttggc caaccagcag gtgtttcaca tcagctgctt 1380
ccgttgctcc tattgcaaca acaaactcag tctaggaaca tatgcatctt tacatggaag 1440
aatctattgt aagcctcact tcaatcaact ctttaaatct aagggcaact atgatgaagg 1500
ctttgggcac agaccacaca aggatctatg ggcaagcaaa aatgaaaacg aagagatttt 1560
ggagagacca gcccagcttg caaatgcaag ggagacccct cacagcccag gggtagaaga 1620
tgcccctatt gctaaggtgg gtgtcctggc tgcaagtatg gaagccaagg cctcctctca 1680
gcaggagaag gaagacaagc cagctgaaac caagaagctg aggatcgcct ggccaccccc 1740
cactgaactt ggaagttcag gaagtgcctt ggaggaaggg atcaaaatgt caaagcccaa 1800
atggcctcct gaagacgaaa tcagcaagcc cgaagttcct gaggatgtcg atctagatct 1860
gaagaagcta agacgatctt cttcactgaa ggaaagaagc cgcccattca ctgtagcagc 1920
ttcatttcaa agcacctctg tcaagagccc aaaaactgtg tccccaccta tcaggaaagg 1980
ctggagcatg tcagagcaga gtgaagagtc tgtgggtgga agagttgcag aaaggaaaca 2040
agtggaaaat gccaaggctt ctaagaagaa tgggaatgtg ggaaaaacaa cctggcaaaa 2100
caaagaatct aaaggagaga cagggaagag aagtaaggaa ggtcatagtt tggagatgga 2160
gaatgagaat cttgtagaaa atggtgcaga ctccgatgaa gatgataaca gcttcctcaa 2220
acaacaatct ccacaagaac ccaagtctct gaattggtcg agttttgtag acaacacctt 2280
tgctgaagaa ttcactactc agaatcagaa atcccaggat gtggaactct gggagggaga 2340
agtggtcaaa gagctctctg tggaagaaca gataaagaga aatcggtatt atgatgagga 2400
tgaggatgaa gagtgacaaa ttgcaatgat gctgggcctt aaattcatgt tagtgttagc 2460
gagccactgc cctttgtcaa aatgtgatgc acataagcag gtatcccagc atgaaatgta 2520
atttacttgg aagtaacttt ggaaaagaat tccttcttaa aatcaaaaac aaaacaaaaa 2580
aacacaaaaa acacattcta aatactagag ataactttac ttaaattctt cattttagca 2640
gtgatgatat gcataagtgc tgtaaggctt gtaactgggg aaatattcca cctgataata 2700
gcccagattc tactgtattc ccaaaaggca atattaaggt agatagatga ttagtagtat 2760
attgttacac actattttgg aattagagaa catacagaag gaatttaggg gcttaaacat 2820
tacgactgaa tgcactttag tataaagggc acagtttgta tatttttaaa tgaataccaa 2880
tttaattttt tagtatttac ctgttaagag attatttagt ctttaaattt tttaggttaa 2940
ttttcttgct gtgatatata tgaggaattt actactttat gtcctgctct ctaaactaca 3000
tcctgaactc gacgtcctga ggtataatac aacagagcac tttttgaggc aattgaaaaa 3060
ccaacctaca ctcttcggtg cttagagaga tctgctgtct cccaaataag cttttgtatc 3120
tgccagtgaa tttactgtac tccaaatgat tgctttcttt tctggtgata tctgtgcttc 3180
tcataattac tgaaagctgc aatattttag taataccttc gggatcactg tcccccatct 3240
tccgtgttag agcaaagtga agagtttaaa ggaggaagaa gaaagaactg tcttacacca 3300
cttgagctca gacctctaaa ccctgtattt cccttatgat gtcccctttt tgagacacta 3360
atttttaaat acttactagc tctgaaatat attgattttt atcacagtat tctcagggtg 3420
aaattaaacc aactataggc ctttttcttg ggatgatttt ctagtcttaa ggtttgggga 3480
cattataaac ttgagtacat ttgttgtaca cagttgatat tccaaattgt atggatggga 3540
gggagaggtg tcttaagctg taggcttttc tttgtactgc atttatagag atttagcttt 3600
aatatttttt agagatgtaa aacattctgc tttcttagtc ttacctagtc tgaaacattt 3660
ttattcaata aagattttaa ttaaaatttg aaaaaaaaaa aaaaa 3705

<210> 17
<211> 1273
<212> DNA
<213> Homo sapiens

<n>
<220>
<221> unsure
<222> 1237
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte ID No: 2768356CB1

<400> 17
gccaccatcc gcctatctct gcgtgtcatg ctgagtctag aaattttgtt ttctggcaag 60
atgtgagatg gaaaaacaaa ttctggggca aatccatgga aattgttcca attggcacaa 120
cccatgtgac tctgccagtt tttggggatc attttgagtg gaacaaagtg acctcttgca 180
tccataacat cttaagcggg cagaggtgga ttgagcacta tggagagatt gtcatcaaga 240
acctgcatga tgattcctgc tactgcaaag tgaattttat aaaggcaaaa tactggagca 300
ctaatgccca tgagattgaa ggcacagtgt ttgacaggag tggaaaagcg gttcatcggc 360
tgtttgggaa atggcatgaa agcatctact gtggcggcgg ctcctcttct gcctgtgtat 420
ggagagcaaa tcctatgccg aaaggctacg agcaatacta tagcttcaca cagtttgcgc 480
tggaattaaa tgaaatggat ccatcatcaa agtctttatt gccacctact gacactcgat 540
ttaggccaga ccagaggttt ctagaggaag ggaacttaga agaagctgaa atacaaaagc 600
agaggattga acaactgcag agagaaaggc ggcgggtctt agaagaaaat catgtggagc 660
accagcctcg gtttttcagg aaatccgacg atgactcttg ggtgagcaac ggcacctatt 720
tggaacttag aaaagatctt ggtttttcca aactggacca tcctgtctta tggtgaaaaa 780
gtaaagaaga aagataacat tagtgtattt ctcctgtgct tgccttctga agtggcacaa 840
acctgtgttt atatatttaa aagatactct aggatgatca cttgtgctta gcttagcatt 900
gtaactcttt aagtctatat tttcctcagt gcgtttcttt acaatttcaa atgttaccct 960
gattgtttat atgaatgtag aacaccttga catttctttt tatatataaa ctatttaata 1020
aaaatgaaag attgaatgtt catgtgtggg ttaaaaaaag aagctttaac actaattttc 1080
caaaggttag ggaagattcc aattaaattt atgccttata aaattatgtt gtagaaaaaa 1140
tatcaacctc tcccaggtgc attaagaaat aagaattccc agggttactc acccatgcgt 1200
aagctaccaa gtttaatttg gtagctgaaa tatcttnttg cctcagacag ctcttgaatt 1260
gctcatacag aca 1273

<210> 18
<211> 1733
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5324145CB1

<400> 18
gccaacattc taggattctg ctggactagt tcaactgaaa ttgtcttcat aacagatcaa 60
ggaatcgaat tttaccaggt attaccagag aaacggagtc tgaaactctt gaagagccac 120
aatctcaatg tgaattggta catgtactgc cccgagagcg ccgtgatctt gctgtctacc 180
acggtcctgg agaatgtcct gcagcctttt cactttaggg ctggcactat gtcgaagctg 240
cccaaatttg agattgaatt accagctgcg cctaagtcaa ctaaacccag cctttccgaa 300
agagacatcg caatggctac catatacggg cagctgtatg ttctcttctt gaggcatcat 360
tctcggacct ccaacagcac aggagcggag gtggtcctct atcatctacc acgagaaggt 420
gcctgtaaaa agatgcacat attgaagtta aataggacgg gaaagtttgc cctgaacgtg 480
gtggacaacc tggtagtcgt gcatcatcag gatacagaga catcggtaat attcgatatc 540
aagttacggg gagagtttga cggctccgtt accttccacc accccgtgct tcctgctcga 600
tcgatccagc cctatcagat ccccatcaca ggtcctgctg ccgtgaccag ccagtctcct 660
gttccatgta aactctattc ttcatcttgg attgtctttc aacctgacat cattatcagc 720
gcaagccaag gttacctctg gaacctccaa gtgaaacttg agcccatagt aaatctctta 780
ccagacaaag gaagactcat ggactttctc ctccagagaa aggaatgcaa gatggtcatc 840
ctgtctgtct gttcacagat gttaagtgag tcagacagag catcgctgcc cgtgatagcc 900
actgtttttg ataaactcaa ccatgagtat aaaaagtacc tggatgccga gcagagttat 960
gcgatggcgg tggaagcagg gcagagccga agcagcccgc tcctcaagag gccggtgcgg 1020
acccaggcgg tgctggacca gtcagatgtg tacacccatg tcctgtcagc ctttgtggaa 1080
aagaaggaga tgcctcataa atttgtgata gccgtgctga tggaatacat tcgttctctt 1140
aaccagtttc agattgcagt acagcattac ctacatgaac ttgttatcaa aacccttgtc 1200
cagcacaacc tcttttatat gctgcatcag ttcctgcagt accacgtcct cagcgactcc 1260
aaaccttttgg cttgtctgct gttatcccta gagagtttct atcctcctgc tcatcagcta 1320
tctctggaca tgctgaagcg actttcaaca gcaaatgatg aaatagtaga agttctcctt 1380
tccaaacacc aagtgttagc tgccttaagg tttatccggg gcattggtgg ccatgacaac 1440
atttctgcac gaaaattttt agatgctgca aagcagactg aagacaacat gcttttctat 1500
acaatattcc gctttttttga acagcgaaac cagcgtttgc gagggagccc caatttcaca 1560
ccaggggaac actgtgaaga acatgttgct tttttcaaac agatttttgg agaccaagct 1620
ctaatgaggc ctacaacatt ctgaaatcac ttgctgtttt tttatataaa aatgtgtaca 1680
aagttaattt attgcattaa taaagctctt taaactataa aaaaaaaaaa aaa 1733

<210> 19
<211> 1370

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1004646CB1

<400> 19
gcacagcctc gcaatgagct gccacaactg ctccgacccc caggtccttt gcagctccgg 60
gcagctgttc ctgcagcccc tctgggacca cctgaggagc tgggaggccc tcctacagtc 120
gcccttcttc ccggtcatct tctccatcac cacatacgtg ggcttttgcc tgcccttcgt 180
ggtcctggat atcctgtgct cctgggtgcc cgccctgcgg cgctacaaga tccaccctga 240
cttctcgcca tccgcgcagc agctgctacc ttgcctgggg cagaccctct accagcatgt 300
gatgtttgtg ttccccgtga cgctgctgca ttgggcccgc agcccggccc tcctgcccca 360
cgaagctccc gagctgctcc tgctgctgca ccacatcctg ttctgcctgc tactcttcga 420
catggagttc ttcgtgtggc acctgctgca ccacaaggtg ccctggctgt accgcacctt 480
ccacaaggtg caccaccaga actcgtcctc gttcgcgctg gcaacgcagt atatgagcgt 540
ctgggaactg ttttctttgg gcttcttcga catgatgaac gtcacactgc tcgggtgcca 600
cccgctcacc accctgacct tccacgtggt caacatctgg ctttccgtgg aggaccactc 660
cggctacaac ttcccttggt ccactcacag actggtgccc ttcgggtggt acgggggtgt 720
ggtgcaccac gacctgcatc actctcactt taactgcaac ttcgctccgt actttacaca 780
ctgggacaaa atactgggaa cgctgcggac tgcatctgtc ccagcgcggt gatgtggctg 840
cggtgggtgc ccctaagact cgggactgct gtgcctttca cacttgaatg aagagaaaca 900
cctgagctat atatttttt aaagcaacta acttattgct ttatgtttat ctatgaaaac 960
catagataaa atctgatgca tttttgtaat ctgacaaagt aatttacata ctgtttgtgt 1020
atcaatacaa ttttgtgttc ttggtattct tagtctagct cacctcaata gccttgaatc 1080
ctgcatatga attagacatt catcactggc atatttagaa tatctctaaa aggacttgtt 1140
tgtagaataa ggaattttct atgtttcaaa gtgttctaaa acctggctaa aagaatgtat 1200
ttttgtggat ggtgttgact tctgactcta aaagcaatca aacatgtttc tgctggacag 1260
tgaccaagaa ttatagtacc ttcttatatt tttttataga actgtatatt tattttgaaa 1320
gaaatgttat tcgtgcttta aaaaggaaaa aaaaccatga atcaaataaa 1370

<210> 20
<211> 1264
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1802851CB1

<400> 20
ctcgcgtgcg gtgcgcaggg ataagagagc ggtctggaca gcgcgtggcc ggcgccgctg 60
tggggacagc atgagcggcg gttggatggc gcaggttgga gcgtggcgaa caggggctct 120
gggcctggcg ctgctgctgc tgctcggcct cggactaggc ctggaggccg ccgcgagccc 180
gctttccacc ccgacctctg cccaggccgc aggccccagc tcaggctcgt gcccacccac 240
caagttccag tgccgcacca gtggcttatg cgtgcccctc acctggcgct gcgacaggga 300
cttggactgc agcgatggca gcgatgagga ggagtgcagg attgagccat gtacccagaa 360
agggcaatgc ccaccgcccc ctggcctccc ctgcccctgc accggcgtca gtgactgctc 420
tgggggaact gacaagaaac tgcgcaactg cagccgcctg gcctgcctag caggcgagct 480
ccgttgcacg ctgagcgatg actgcattcc actcacgtgg cgctgcgacg gccacccaga 540
ctgtcccgac tccagcgacg agctcggctg tggaaccaat gagatcctcc cggaagggga 600
tgccacaacc atggggcccc ctgtgaccct ggagagtgtc acctctctca ggaatgccac 660
aaccatgggg cccctgtga ccctggagag tgtcccctct gtcgggaatg ccacatcctc 720
ctctgccgga gaccagtctg gaagcccaac tgcctatggg gttattgcag ctgctgcggt 780
gctcagtgca agcctggtca ccgccaccct cctccttttg tcctggctcc gagcccagga 840
gcgcctccgc ccactggggt tactggtggc catgaaggag tccctgctgc tgtcagaaca 900
gaagacctcg ctgccctgag gacaagcact tgccaccacc gtcactcagc cctgggcgta 960
gccggacagg aggagagcag tgatgcggat gggtacccgg gcacaccagc cctcagagac 1020
ctgagctctt ctggccacgt ggaacctcga acccgagctc ctgcagaagt ggccctggag 1080
attgagggtc cctggacact ccctatggag atccggggag ctaggatggg gaacctgcca 1140
cagccagaac tgaggggctg gccccaggca gctcccaggg ggtagaacgg ccctgtgctt 1200
aagacactcc tgctgccccg tctgagggtg gcgattaaag ttgcttcaca tcctcaaaaa 1260
aaaa 1264

<210> 21

<211> 2566
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2764333CB1

<400> 21
cccacgcgtc cgacgagttc tatgatgcgc tgtcagattc cgagtccgaa aggtccctga 60
gtagattgga agcagtgaca gcacgctcct ttgaagagga aggagagcat ttgggcagta 120
gaaaacacag aatgtccgaa gaaaaagact gtggtggcgg agatgctctc tccaatggca 180
tcaagaaaca cagaacaagt ttgccttctc ctatgttttc cagaaatgac ttcagtatct 240
ggagcatcct cagaaaatgt attggaatgg aactatccaa gatcacgatg ccagttatat 300
ttaatgagcc tctgagcttc ctacagcgcc taactgaata catggagcat acttacctca 360
tccacaaggc cagttcactc tctgatcctg tggaaaggat gcagtgtgta gctgcgtttg 420
ctgtatctgc tgttgcttct cagtgggaac ggactggaaa acctttcaac ccactgctgg 480
gagagactta tgaattagtg cgagatgacc ttggatttag actcatctcc gaacaggtca 540
gccatcaccc accaatcagt gcatttcatg ctgaaggatt aaacaatgac ttcatctttc 600
atggctctat ctatcccaaa ctgaaattct gggggaagag tgtagaagca gaacccaaag 660
gaaccatcac cttggagctc cttgaacaca atgaggcata tacatggaca aatcccacct 720
gctgtgtgca taatatcatt gtgggtaaac tgtggatcga acagtatggc aatgtggaaa 780
ttataaacca caagactggg gacaaatgtg tgttgaattt taagccatgt ggccttttg 840
gtaaggaatt acacaaagtt gaaggctaca ttcaagataa aagcaaaaag aagctctgtg 900
ccctctatgg gaagtggact gaatgtttat acagtgttga ccctgccacg tttgacgctt 960
acaaaaaaaa tgataagaaa aatacagaag agaagaagaa cagcaaacag atgagcacct 1020
ctgaggagtt ggatgaaatg ccagtgccgg attctgaaag tgtattcatt atccctggaa 1080
gcgttcttct atggcgaata gccccacggc ctccaaattc tgcccagatg tataatttta 1140
ctagttttgc aatggttttg aatgaagtag acaaagacat ggagagtgtg attcccaaga 1200
cagactgcag gttacggcct gacatcagag ccatggaaaa tggagagata gatcaagcta 1260
gtgaagaaaa aaaacgactt gaggaaaaac aaagagcagc ccgcaaaaac aggtccaagt 1320
cagaagagga ctggaagacg aggtggttcc atcaaggtcc taatccctac aatggagcac 1380
aggactggat ttactctggc agctactggg acagaaatta cttcaatttg cctgacattt 1440
attaaaatgc atacaagtca gggtgtttgg ctaatctaca aataagtctt aaacctatgt 1500
ttttaaattt ttttcccttg gtttctactt atcttttaaa aaaaaaaatg aaaaaacact 1560
catgagataa ctgcatttca cccaacaaaa gcagggtata aggcgatatt ggtgatgaaa 1620
gtcttaggaa aaatgcataa ttttgctata aaatgtactt atttggaata ctattttata 1680
tagaggtaag agaacactgc tggggaatat gctttttatg gttgctgttg ccatatttac 1740
tgaaggttta tacctaaatg taactttagc tttatggaac tatatagtaa tcccaaatca 1800
agttatttttg aatatttttta tgctgtcatg cttgaatgtt ttagatgtaa cctttgacat 1860
atttagaact ctcctcctat acaatgttta ttctcagata tagaggttat gtcatttat 1920
aaagacttca ttgataagat ggcttttatt catactaatc ctcccaatgt tacccccttcc 1980
atcttccaag aagaaaaaaa atgcctgaat attcagaata gatatttctg atttgaaaat 2040
tctaaagaat taaactggaa aagtatttca tttacttagt gctctgaatt tacttttaca 2100
gttttctgca gtcagtatca ttaaaatggt taagtttaca tttgaactga aaatatgtat 2160
aaaatctagc aattcacaaa aatgccctag aaatatagat tttaatcacc attacataat 2220
gacaaacctt gttaaatgct tccacttcca gtggcaaatg ccactaggga aagtaagttg 2280
cactcatgta agtatcaaac tatataaaag gaggccttgt gcatttcaag tttgcaaagt 2340
acctgtgtac ttaaaatatg tgtggagacc tactgtacag tagttttgcc cctttaattg 2400
gggcacattc atcttaaatc ttatagtatt tatccaccca aaccccagac tgagatactg 2460
ctcccagggg cctaggtagc tgccagtccg tgattttaat tgctgtcttg aagttaacaa 2520
gtgttataat gaaataatct acctgatgct aaataaaggc tttaga 2566

<210> 22
<211> 1848
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2798021CB1

<400> 22
aaccagctga aaagcatgcg ccaagctgca gcagatgcaa agcctgaaag tttaatgaag 60
aggctagagg aggagataaa atttaattta tatatggtaa ctgaaaaatt tcctaaagaa 120
ttagaaaata agaaaaagga attacatttt ttacaaaaag tagtttcaga gccagctatg 180
ggccattctg atcttcttga acttgaatct aaaataaatg aaataaacac agaaattaac 240

cagttgattg aaaagaaaat gatgagaaat gagcccattg aaggcaaact ctcactgtat 300
aggcaacagg catctatcat ttcccgtaaa aaagaagcca aagctgagga acttcaggag 360
gccaaggaga agttagccag cctagagaga gaagcatcag taaagagaaa tcagacccgt 420
gaatttgatg gtactgaagt tttaaaggga gatgagttca aacgatatgt caataaactt 480
cgaagcaaga gtacagtttt caaaaagaag catcagataa tagctgaact taaagctgaa 540
ttcggtcttt tgcagaggac tgaagaactt cttaagcaac gtcatgaaaa tattcaacaa 600
caactgcaaa ctatggagga gaaaaagggt atatctggat atagttacac ccaagaagag 660
ctagaaagag tatctgcact gaagagtgaa gttgatgaaa tgaaaggacg aacattggat 720
gatatgtctg aaatggtgaa aaaactgtat tcattggtat ctgaaaagaa gtcagctctt 780
gcctcagtta taaaagagct acgacagttg cgtcaaaaat atcaagaact gacccaggag 840
tgtgatgaaa agaaatccca gtatgatagc tgtgcagcag gcctcgaaag caatcggtcc 900
aaattagaac aggaagttag aagactccgt gaagaatgtc ttcaagaaga aagtagatac 960
cattatacaa attgtatgat taagaaccta gaagttcaac ttcgtcgtgc tactgatgag 1020
atgaaggcat atatctcttc tgatcaacaa gaaaaaagaa aggcaattag ggaacagtat 1080
accaaaaata ctgctgaaca agaaaacctt ggaaagaaac ttcgggaaaa acaaaaagtt 1140
atacgagaaa gtcatggtcc aaatatgaaa caagcaaaaa tgtggcgtga tttggaacaa 1200
ttaatggaat gtaagaaaca gtgctttctg aaacaacaaa gccaaacttc cattggtcag 1260
gtaattcagg agggtgggga ggaccggcta atactgtgaa ttcttgtgtc atcgtttggg 1320
gttttacttg ataccactag ctataagcct aatctcataa tgtatttctt ttttgaaact 1380
gatttgtata gcatttttgtt ttcagaagag ccattcttta ttaagtttc atagaaaata 1440
atgttaaggt agatttagtt tgaatgtttt ttcatatgaa aaagaggctt ttattctttt 1500
ccatagttta gacatcactg gcgtcttctg agttttatga gacaggaaac taagtttact 1560
atctgtaaat gtaaacatat gtccattaag aaacatgtag ttttttttta gaatgtaata 1620
acccagtggc ttactgtttt tcttaatctc ttttaaaaaa actttagaag aatcttttag 1680
gaactaatat ctcttgttct gaagaaacat ttatctgacg ttcagcagtt cctacagttt 1740
tacttcagtt tatttttctt ctgtaaaatg caagaaaatt taatattttg actaacatgt 1800
cttttctgtt tgtatcattt aaaggcatat aaacttgtcg agtattaa 1848

<210> 23
<211> 2518
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3335404CB1

<400> 23
cgcagagccg gcgagggctt gcggtcctgg acggtagggg tctgcgtctg gtgccacctt 60
ctccatcagg ctgcttgctg ggtccaccaa gcctgacccc tccagccagg aggggatcac 120
ggagtgtgtc ccccgcccgg gctgctgcct gccggaaggg ggctgggaaa ccgggatccc 180
cgcgcgccac ttgcccggca gcgtcagttt cacccaggac tggctagcgg ttccctcgtg 240
gctctgcgcg gaggtccgcc tcctcccttc cctcccccat cccgcgcgct gcctccgcct 300
ccttcttctg cctccaccct caacccccat ccccgcctga cgggagctag ccctcagtcc 360
gcccgagctg tggttgtggg cgccggacaa gtccaaggcg cctcctccca atatggacag 420
ccgctacaac agcactgcgg gcatcgggga cttgaaccag ctgagcgctg ccatcccggc 480
cacgcgggtg gaggtgtccg tgtcctgcag aaatcttctt gacagagaca cattttctaa 540
atctgatcca atttgtgtct tatatgtaca aggagttgga aataaagaat ggagagagtt 600
tggaagaact gaagtaattg ataatacttt aaatcctgat tttgtaagaa agtttattct 660
ggactacttt tttgaagaaa gagagaatct tcgttttgac ttgtatgatg ttgattcaaa 720
gagccccaac ttatccaaac atgactttct gggacaagtg ttttgtacat tgggagagat 780
cgttggttca cagggaagtc gcctggaaaa accaatagta ggaattccag ggaagaaatg 840
tggtacaatc atacttacag cagaggaatt aaactgttgc agggatgccg ttttgatgca 900
attttgtgcg aacaaattgg acaagaagga cttctttgga aaatcagatc ctttccttgt 960
attttatcga agtaatgaag atggcagttt tacaatttgt cacaagacag aagttgtcaa 1020
aaacactcta aatccagtat ggcaagcatt caagatctca gtcagagcat tatgtaatgg 1080
agactatgac agaacaatca aagtagaggt gtatgactgg gaccgagatg gaagtcatga 1140
tttcattgga gaatttacaa caagctatag ggaacttttct agagggcagt cacaattcaa 1200
cgtatatgag gtggtgaatc ccaaaaagaa aggaaaaaag aaaaaatata ctaattctgg 1260
aacagtaact ttactctctt tcttggtaga aacagaagtt tcattccttg actacattaa 1320
gggagggacg caaatcaatt tcacagtggc tattgatttt acagcatcaa acgggaaccc 1380
tgctcagccc acttccctcc actacatgaa tccttaccaa ctgaatgcct atggtatggc 1440
actaaaagca gtgggagaaa ttgttcaaga ttatgacagt gataaaatgt ttccagctct 1500
aggatttggt gcaaaactgc ctccagatgg aaggatatct cacgaatttg ctttgaatgg 1560
gaatcctcaa aacccctact gtgatggcat tgaggggggtc atggaggctt attacaggag 1620
tctgaaatct gtacaactat atgggcccac caactttgct cctgtaatta atcatgtagc 1680
aagatatgct tcttctgtaa aggatggctc ccagtatttt gtgcttctga ttgttacaga 1740

tggtgttatc tcagatatgg cccagactaa ggagtccata gttaatgcct caaaacttcc 1800
aatgtcaata attatagtag gtgttggacc agcagaattt gatgcaatgg tcgaattgga 1860
tggagatgat gtaagagtct cctctagagg aaaatatgct gaaagagaca ttgtgcagtt 1920
tgtgccattc agggattata ttgacagaag tggaaaccac atactgagca tggctagatt 1980
ggctaaagat gtcctagctg agatccctga gcagtttctc tcctatatga gagcccgagg 2040
aatcaagcca tcacctgcgc ctcccccata caccccacct acacatgtgt tacagactca 2100
aatatgactg tgctctgaaa tgctaatgtc aactacaaat caaaagtgct gagttaatgc 2160
tttgtgcctg gtgctctgta atgaaccagg caatgagata gttttctcag tttggtttca 2220
gcagttaatg tgctttcttg gatccaaatt taaatatctt cctaaaccaa aactgtaaat 2280
atggttgttg catgagcaac agaaaaaatt gtttaaatgc ttgaagcaaa gtatggatgt 2340
cttctctaaa tctttctttc ttttttttct ttttttttct ttaacagaaa cagctaattt 2400
ccaatgtatt gttgggaaaa agcacaaact gtgtttttaa ctcaaatatt gtcttcgact 2460
gctatgtgta taggaaagca gtctctgtat caatgtttac atgttactac tttttaaa 2518

<210> 24
<211> 1120
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3735780CB1

<400> 24
gaaatccagt tatcaaaatt gactcaagaa gagagaacct aacagaacaa taacaatgga 60
agaaattggg aacattatca caaagctatc atcctgccaa actccaggct cagatgtcac 120
aggttaaaaa aagtccttca tgaaaaagaa agatcttaag cagcatgatg gattcagaag 180
ctcatgaaaa gaggccacca atactaacat cttcaaaaca agatatatca cctcatatta 240
caaatgttgg tgagatgaag cattacttgt gtggctgctg tgcagccttc aacaatgtcg 300
caatcacatt tcccattcag aaggtcctct ttcgacaaca gctgtatggc atcaaaaccc 360
gggatgcaat acttcagttg agaagggatg gatttcgaaa tttgtatcgt ggaatccttc 420
ccccattgat gcagaagaca actacgcttg cacttatgtt tggtctgtat gaggatttat 480
cctgccttct ccacaagcat gtcagtgctc cagagtttgc aaccagtggc gtggcggcag 540
tgcttgcagg gacaacagaa gcaattttca ctccactgga aagagttcag acattgcttc 600
aagaccacaa acatcatgac aaatttacca acacttacca ggctttcaag gcactgaaat 660
gtcatggaat tggagagtat tatcgaggct tggtgcccat tcttttccgg aatggactca 720
gcaatgtctt gtttttcggc cttcgaggtc ccattaagga gcatctgcct accgcaacga 780
ctcacagtgc tcatctggtc aatgatttta tctgtggagg tctattgggt gccatgttgg 840
gattcttgtt ttttccaatt aatgttgtaa aaactcgcat acagtctcag attggtgggg 900
aatttcagtc tttccccaag gttttccaaa aaatctggct ggaacgggac agaaaactga 960
taaatctttt cagaggtgcc catctgaatt accatcggtc cctcatctct tggggcataa 1020
tcaatgcaac ttatgagttc ttgttaaagg ttatatgaaa aaaccatcag ttaagtgcca 1080
tttatcaact gaatagacct tctaagaaga aaaaaaaaaa 1120